

Dynamikkompression in der digitalen Verkehrsphotographie durch nichtlineare Filterung in einer Laplacepyramide

Diplomarbeit
im Fachbereich Photoingenieurwesen und Medientechnik
an der Fachhochschule Köln

Autor
Danny Odenthal
Leverkusen
Mat.-Nr. : 11020445

Referent : Prof. Dr. D. Kunz
Köln
Korreferent : Dipl.-Ing. A. Behrens, Robot Visual Systems
Monheim a.R.

Köln, Dezember 2002

Dynamic range reduction
in digital traffic
photography
by non-linear filtering
in a Laplacain Pyramid

Thesis
at the Department
of
Imaging Sciences and Media Technology
Univesity of Applied Sciences Cologne

Author
Danny Odenthal
Leverkusen
Mat.-Nr. : 11020445

First Reviewer : Prof. Dr. D. Kunz
Cologne
Second Rewiewer : Dipl.-Ing. A. Behrens, Robot Visual Systems
Monheim a.R.

Cologne, December 2002

Inhaltsverzeichnis

1	Einleitung	5
1.1	Zielsetzung	5
1.1.1	Dynamikkompression	5
1.1.2	Überschwinger	6
1.1.3	Lokaler Kontrast	6
1.1.4	Rauschreduktion	6
1.2	Kurze Darstellung der wesentlichen Grundlagen	7
1.2.1	Digitale Bilder allgemein	7
1.2.2	Histogramm	7
1.2.3	Lineare Kontraststreckung	7
1.2.4	Photometrische Grundlagen	8
1.2.5	Rauschen	9
1.2.6	Abtasttheorem	10
1.2.7	Gausspyramide	11
1.2.8	Laplacepyramide	12
1.2.9	Rekombination	15
1.2.10	Binomialfilter	16
1.2.11	Faltungsoperation	17
1.2.12	Separierbarkeit von Filtern	18
1.2.13	Impulsantwort und Übertragungsfunktion	18
1.3	Bisherige Methoden	19
1.3.1	Zwei 8 bit Bilder	19
1.3.2	Gradationskurven in Photoshop	19
2	Material und Methode	21
2.1	Methode	21
2.2	Software	22
2.3	Testbilder	22
2.4	Darstellung der Bilder	23
2.5	Implementierung pyramide.c	24
2.5.1	Struktur des Programms	24
2.5.2	Ein-/ Ausgabe der Parameter	24

2.5.3	Ablaufdiagramm	25
2.5.4	Reduzierende Faltung	26
2.5.5	Expandierende Faltung	27
2.5.6	Logarithmierung	28
2.5.7	Filterung in den Laplaceebenen	29
2.5.8	Filterkern	30
3	Testergebnisse	31
3.1	Logarithmierung	31
3.2	Wahl der Anzahl der Laplacestufen	32
3.3	Lineare Filterung	35
3.4	Nichtlineare Filterung	37
4	Diskussion	44
4.1	Zusammenfassung	47
5	Anhang	48
5.1	Literaturverzeichnis	48
5.2	Quellcode pyramide.c	49
5.3	Quellcode der Funktion „Auslesen_8_bit“	57
5.4	Quellcode zeile.c	58
5.5	Verwendete Bildformate	60
5.5	Testbilder	61
5.6	Eidesstattliche Erklärung	72
5.7	Sperrvermerk	72
5.8	Weitergabeerklärung	72

1 Einleitung

1.1 Zielsetzung

In der Verkehrsphotographie werden Aufnahmen von Verkehrsteilnehmern gemacht, um Ordnungswidrigkeiten, wie zu schnelles Fahren oder Überqueren einer Ampel bei Rotlicht, Fahrzeugen und Personen zuordnen zu können. Das Problem bei diesen Aufnahmen liegt in der hohen Dynamik, die vor allem durch Reflexionen im Nummernschildbereich entstehen. Zur Zeit wird dieses Problem dadurch gelöst, indem man aus der Originalaufnahme zwei Bilder macht. Eins für den dunklen Bereich, in dem der Fahrer zu erkennen ist und ein Bild für den hellen Bereich, in dem das Nummernschild erscheint (→ 1.3.1).

Ziel dieser Arbeit ist es, die Dynamik so weit zu komprimieren, dass die wesentlichen Elemente, wie das Gesicht, das Kennzeichen und die Skriptzeile, die alle Informationen über die Aufnahme enthält, in einem Bild zu erkennen sind, anstatt wie oben beschrieben in zwei Bildern. Dadurch kann der Speicheraufwand für die Bilder um die Hälfte reduziert werden.

Im Speziellen muss darauf geachtet werden, dass störende Artefakte, die bei den gewählten Methoden (→ 2.1) entstehen können, vermieden werden. Dies sind besonders Überschwinger an scharfen Kanten, die z.B. an den Buchstaben des Kennzeichens entstehen und niedrige lokale Kontraste im Bereich des Fahrers.

Zusätzlich wird auf das Problem des Rauschens eingegangen. Durch dessen Reduktion kann der subjektive Eindruck des Bildes stark verbessert werden.

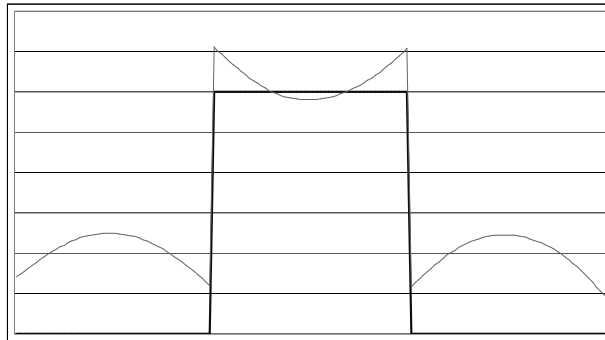
In dieser Arbeit werden nur Schwarz/Weiß Bilder behandelt, weil es das Standardaufnahmeverfahren bei der Verkehrsüberwachung innerhalb der Bundesrepublik Deutschland ist. In anderen Ländern werden zwar auch Farbbilder aufgenommen, dessen Behandlung aber nicht Bestandteil dieser Arbeit sein soll.

1.1.1 Dynamikkompression

Die Bilder liegen im 16-bit Format vor, d.h. es können max. bis zu 65.536 Graustufen dargestellt werden, welche aber weder von der digitalen noch von der analogen Aufnahme voll ausgenutzt werden. Die digitalen CCD Chips sind in der Dynamikfrage dem analogen Nassfilm stark unterlegen. Während ein Nassfilm bis zu 14 bit Informationen enthält, besitzt ein CCD Chip maximal 11 bit. Für eine jpg. Komprimierung muss das Bild im 8 bit Format vorliegen, d.h. es dürfen max. 256 Graustufen vorhanden sein, die alle relevanten Elemente des Bildes aufweisen.

1.1.2 Überschwinger

Durch die gewählten Methoden, die Dynamik zu komprimieren, können an scharfen Kanten Überschwinger entstehen, da sie im wesentlichen auf einem Hochpassfilter beruhen. Dieses Phänomen stört den harmonischen Eindruck des Bildes und muss deshalb möglichst vermieden werden.



*Abb. 1.1.1 : Überschwinger einer Hochpassfilterung
dicke Linie : Original Kante; dünne Linie : Hochpassbild*

1.1.3 Lokaler Kontrast

Hohe lokale Kontraste, vor allem im Bereich des Fahrers und des Kennzeichens, sind sehr wichtig, weil sie später als Beweismittel dienen sollen. So können Ordnungswidrigkeiten im Verkehr einem Fahrzeugfahrer eindeutig zugeordnet werden. Durch die gewählten Methoden der Dynamikkompression kann es passieren, dass diese Kontraste und damit die Erkennbarkeit stark absinkt. Deshalb ist es nötig, die Methoden so zu modifizieren, dass ein geeigneter Kompromiß zwischen Dynamikkompression und einem hohen lokalen Kontrast gefunden wird.

Ein hoher Kontrast im Nummernschild ist auch deshalb wünschenswert, weil die Bilder noch weitere Verarbeitungsschritte, wie z.B. einen automatischen Kennzeichenleser, durchlaufen.

1.1.4 Rauschreduktion

Ein harmonischer Eindruck eines Bildes hängt auch von den überlagerten Störgrößen (→ 1.2.5) ab. Deshalb sollte es ebenfalls Ziel sein, mit geeigneten Modifikationen der Methoden das Rauschen im Bild zu reduzieren.

1.2 Kurze Darstellung der wesentlichen Grundlagen

1.2.1 Digitale Bilder allgemein

Ein digitales Bild besteht aus einer in physikalische Größen definierten Bildfunktion, die durch eine Bildmatrix repräsentiert wird. Sie wird entweder durch unmittelbare digitale Aufnahme oder durch Digitalisieren eines analogen Bildes erzeugt. Ein farbiges Bild ist durch die drei Bildmatrizen (z.B. die Grundfarben Rot, Grün, Blau) beschrieben [6].

Die Dynamik des Bildes, d.h. der Grauwertumfang, wird in der Einheit „bit“ angegeben:

$$2^{bit} = \text{Anzahl der Graustufen} \quad (1.2.1)$$

1.2.2 Histogramm

Ein Histogramm ist eine graphische Darstellung einer Häufigkeitsverteilung von Elementen, bei der die unabhängige Veränderliche in gleich große Abschnitte (Klassen) eingeteilt ist und die Häufigkeiten darüber aufgetragen sind [7]. Es dient damit zur Visualisierung der Grauwerte-verteilung.

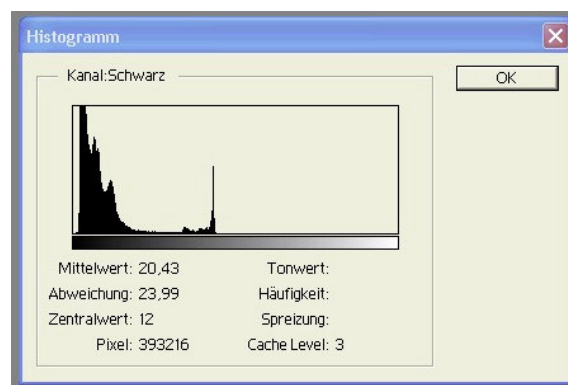


Abb. 1.2.1 : Histogramm

1.2.3 Lineare Kontrastpreizung

Oft befinden sich in einem interessanten Bildbereich die Helligkeitswerte sehr eng beieinander, wohingegen der zur Verfügung stehende Bereich nicht vollkommen ausgeschöpft ist. Deshalb liegt es nahe, den gesamten Bereich auszunutzen. Die dunkelste Stelle wird dabei zu absolutem Schwarz und die hellste zu absolutem Weiß. Die Werte dazwischen werden linear abgebildet. Dieses Verfahren ist sehr einfach, verbessert den Bildeindruck jedoch erheblich [8].

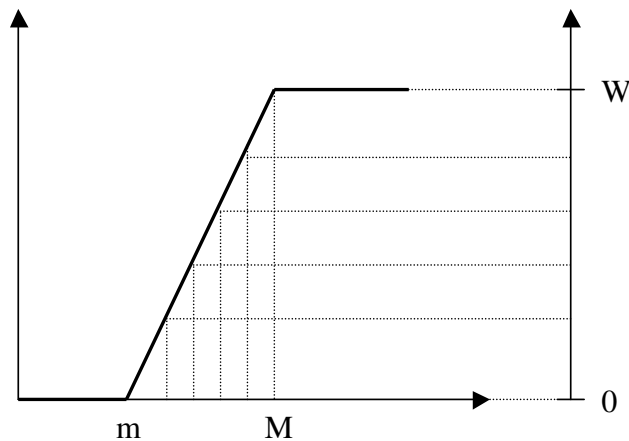


Abb. 1.2.2 : lineare Kontrastspreizung

$$f(x) = \begin{cases} 0 & \text{für } x < m \\ \frac{W}{M-m} (x - m) & \text{für } m \leq x \leq M \\ W & \text{für } x > M \end{cases}$$

(1.2.2)

1.2.4 Photometrische Grundlagen

Eine kleine, punktförmige Lichtquelle, die eine Lichtstärke von $I = 1$ cd (Candela) besitzt, strahlt in einem Raumwinkel von einem sr (Steradian) einen Lichtstrom von 1 lm (Lumen) aus. In 1 m Distanz herrscht dann eine Beleuchtungsstärke von 1 lx (Lux). Ist die Fläche zu 100% rückstrahlend, wird eine Leuchtdichte L von 100% zurückgeworfen. Wirkt die Beleuchtungsstärke von 1 lx während 1 sek ein, so spricht man von einer Belichtung von 1 lxs (Luxsekunde) [1].

Die Leuchtdichte L ist ein Maß einer beleuchteten Fläche. Wird die Fläche als ein Lambert-Strahler betrachtet, d.h. eine Fläche, die in allen Richtungen unter der gleichen Leuchtdichte erscheint, so gilt der Lambert'sche Kosinussatz:

$$L = \frac{I}{A \cdot \cos(\epsilon)}$$

(1.2.3)

mit I = Lichtstärke; ϵ = Winkel der Abstrahlung zum Lot

Für den Remissionsgrad eines Lambert-Strahlers gilt:

$$\rho = \frac{\phi_R}{\phi_o} = \frac{I_o \cdot \pi}{E \cdot A} = \frac{L \cdot A \cdot \pi}{E \cdot A}$$

$$L = \frac{\rho \cdot E}{\pi}$$

(1.2.4)

mit ϕ = Lichtstrom; I = Lichtstärke; E = Beleuchtungsstärke; A = Fläche des Strahlers

Das bedeutet, dass die Leuchtdichte proportional zu dem Produkt aus Remissionsgrad und Beleuchtungsstärke ist.

1.2.5 Rauschen

Bildsignale sind immer von stochastischen Signalschwankungen überlagert:

Ursache hierfür können sein [8] :

- Ungenauigkeit bei der Quantisierung analoger Signale
- Detektorrauschen; jeder Detektor hat ein Dunkelrauschen, abhängig u.a. von der Empfindlichkeit und der Temperatur
- Empfindlichkeitsschwankungen zwischen den Detektoren
- Schwankungen der Beleuchtung
- Quantenmechanische Schwankungen des Lichteinfalls
- Filmkorn bei abgetasteten Analogfotos

Das Rauschen ist in i.d.R. unabhängig vom Ort und von der Bildinformation. Bei digitalen Photos handelt es sich meist um weißes Rauschen, d.h. es ist in allen Frequenzen zum gleichen Anteil vorhanden.

1.2.6 Abtasttheorem

Zur fehlerfreien Rekonstruktion eines Signalverlaufs eines abgetasteten Signals ist eine Mindestabtasthäufigkeit (Abtastfrequenz f_A) erforderlich [9].

Abtasttheorem :

Eine Signalfunktion, die nur Frequenzen im Frequenzband B (bandbegrenztetes Signal) enthält, wobei B gleichzeitig die höchste Signalfrequenz ist, wird durch ihre diskreten Amplitudenwerte im Zeitabstand $t_o = B/2$ vollständig bestimmt.

Andere Formulierung :

Die Abtastfrequenz f_A muss mindestens doppelt so hoch sein wie die höchste im abzutastenden Signal vorkommende Frequenz f_s .

Mit Hilfe des Abtasttheorems kann man sich überlegen, wie groß das Spektrum eines Bildes ist. Dafür braucht man zunächst einmal den Zusammenhang zwischen der Wellenlänge λ und der Frequenz f :

$$f = \frac{1}{\lambda} \quad (1.2.5)$$

Betrachtet man ein Bild der Größe $N \times N$, so ist die kleinste Frequenz durch die Bildgröße N gegeben:

$$\lambda_{max} = N \quad f_{min} = \frac{1}{\lambda_{max}} = \frac{1}{N} \quad (1.2.6)$$

Die kleinste Wellenlänge ist durch das Abtasttheorem begrenzt, da pro Periode mindestens 2 Abtastpunkte benötigt werden:

$$\lambda_{min} = 2 \text{ Pixel} \quad f_{max} = \frac{1}{\lambda_{min}} = \frac{1}{2} \quad (1.2.7)$$

1.2.7 Gausspyramide

Mit „Pyramide“ bezeichnet man eine besondere Bilddatenstruktur, in der Informationen mit unterschiedlicher räumlicher Auflösung verfügbar sind.

Dabei wendet man folgende grundlegende Idee an: Während die Darstellung feiner Skalen die volle Auflösung erfordert, reicht für grobe Strukturen eine niedrige Auflösung. Will man aber die Größe des Bildes reduzieren, kann man nicht einfach abtasten, indem z.B. jeder zweite Bildpunkt in jeder zweiten Bildreihe verwendet wird. Dies würde gegen das Abtasttheorem (\rightarrow 1.2.6) verstoßen und zu Aliasing Effekten führen. Zum Beispiel würde eine Struktur, die im Originalbild dreimal pro Wellenlänge abgetastet wird, im nächsten Bild nur anderthalbmals abgetastet werden. Daraus würde sich ein fehlerhaftes Muster ergeben. Also muss sichergestellt werden, dass alle weniger als viermal pro Wellenlänge abgetasteten Strukturen durch ein passendes Glättungsfilter unterdrückt werden [2].

Die Wiederholung aus kombinierter Glättung und Größenreduktion (\rightarrow 2.5.4) führt zu einer Bilderserie, die aus Tiefpassbildern besteht und die man als Gausspyramide bezeichnet. Die Abbildungen zeigen noch einmal das Prinzip auf:

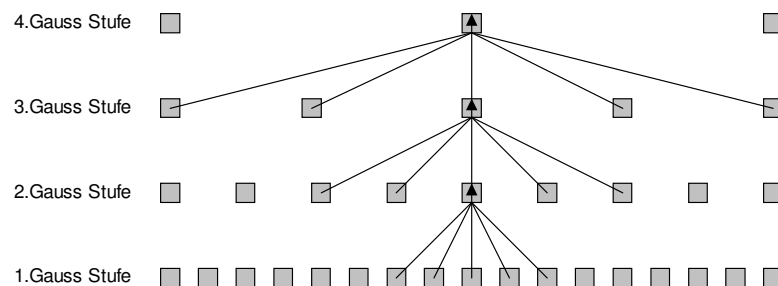


Abb. 1.2.3 : Graphische Veranschaulichung der Erzeugung einer Gausspyramide

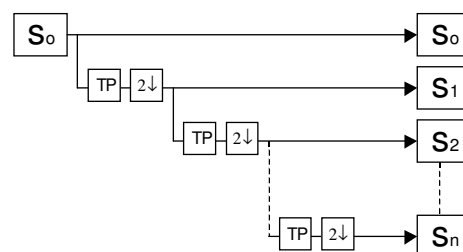


Abb. 1.2.4 : Flussdiagramm einer Gausspyramide

Die Pyramide hat einen sehr geringen Speicherbedarf. Sie benötigt lediglich 1/3 mehr Speicherplatz als das Originalbild.

$$n^2 + \frac{n^2}{4} + \frac{n^2}{16} + \dots = n^2 \left(1 + \frac{1}{4} + \frac{1}{16} + \dots \right) = \frac{n^2}{1 - \frac{1}{4}} = \frac{4}{3} n^2 \quad (1.2.8)$$

Bei einem dreidimensionalen Bild beträgt der Speicherbedarf nur 1/7 mehr. Ebenso effektiv ist die Berechnung der Pyramide. Das selbe Glättungsfilter (\rightarrow 1.2.10) kann auf jeder Ebene angewandt werden und erfordert daher nur 4/3 Rechenoperationen für ein zweidimensionales Bild.



Abb. 1.2.5 : Gausspyramide, nur Darstellung des dunklen Bereichs

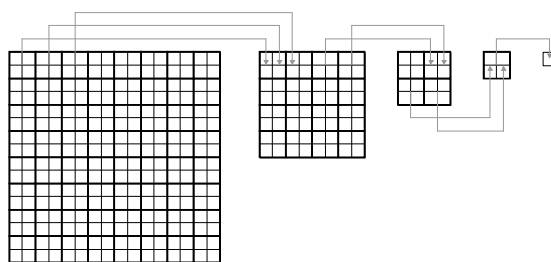


Abb. 1.2.6 : Prinzip zur Erstellung einer Gausspyramide

1.2.8 Laplacepyramide

Aus der Gausspyramide kann man einen weiteren Pyramidentyp berechnen, die sogenannte Laplacepyramide. Sie besteht aus einer Sequenz von Bandpassbildern. Im Gegensatz zur Fouriertransformation (\rightarrow 1.2.13) führt die Laplace Pyramide lediglich zu einer groben Wellenzahlzerlegung ohne Richtungszerlegung. Alle Wellenzahlen innerhalb des Bereichs von

ungefähr einer Oktave (Faktor 2) sind unabhängig von ihrer Richtung in einer Ebene der Pyramide enthalten.

Errechnet wird eine Laplaceebene aus der Differenz zweier Gaussebenen. Dafür muss das Bild der größeren Ebene erst expandiert werden (\rightarrow 2.5.5) :

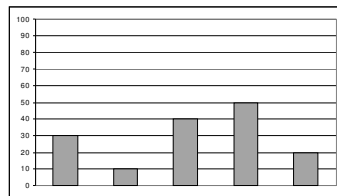


Abb. 1.2.7a; Originalbild

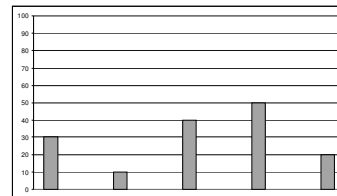


Abb. 1.2.7b; mit Nullen aufgefülltes Bild

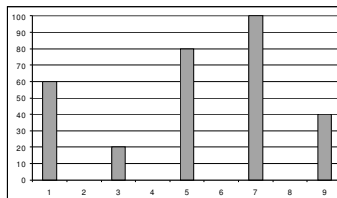


Abb. 1.2.7c; Multiplikation mit 2

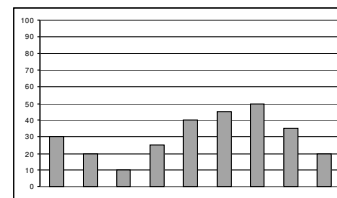


Abb. 1.2.7d; interpoliertes Bild

Abb. 1.2.7 : Interpolation bei der Berechnung einer Laplacepyramide

Die Interpolation wird dadurch erzielt, dass das Bild zuerst mit Nullen aufgefüllt wird (*b*). Anschließend erfährt es eine Multiplikation mit 2 pro Raumrichtung (*c*) und eine glättende Tiefpassfilterung (*d*) [8]. Auch hier kann wieder das gleiche Filter wie bei der Gausspyramide verwendet werden.

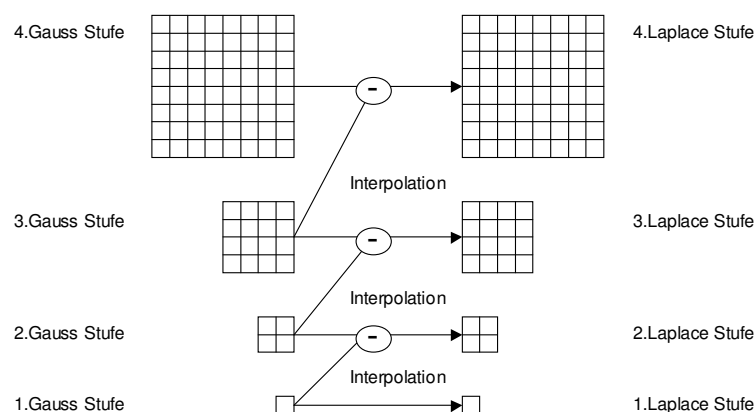


Abb. 1.2.8 : Graphische Darstellung über die Erzeugung einer Laplacepyramide

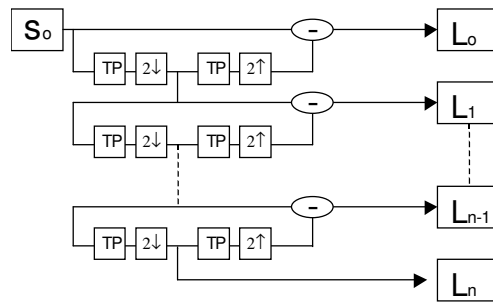


Abb. 1.2.9: Flussdiagramm einer Laplacepyramide

Die Laplacepyramide ist damit ein effektives Schema für eine Bandpasszerlegung. Die zentrale Wellenzahl wird von Ebene zu Ebene halbiert.



Abb. 1.2.10 : Laplacepyramide

Entwickelt wurde die Idee der Laplacepyramide von Burt & Adelson [3] und sollte der effektiven Bildkomprimierung dienen. Bei diesem Verfahren wird nicht das Originalbild komprimiert, sondern ein berechnetes Tiefpassbild und dessen Differenzbild zum Original. Eine Datenkomprimierung tritt deshalb auf, weil das Tiefpassbild mit einer niedrigeren Abtastrate kodiert werden kann als das Originalbild und die Pixel im Differenzbild sehr niedrige Bitzahlen aufweisen. Wiederholt man diesen Vorgang, so repräsentieren die Tiefpassbilder die Gausspyramide und die Differenzbilder die Laplacepyramide. Sehr effektiv ist diese Methode um Bilder progressiv zu übertragen, d.h. der Informationsgehalt eines Bildes wird nicht auf einmal gesendet, sondern anfangend mit dem Tiefpassbild schrittweise aufgebaut.

1.2.9 Rekombination

Bei der Rekombination wird von der untersten Ebene ausgegangen. Zu ihr wird die nächst höhere Ebene dazu addiert. Dabei muss die kleinere Ebene, wie bei der Berechnung der Laplacepyramide aus der Gausspyramide, um den Faktor 2 pro Richtung hochinterpoliert werden (→ Abb. 1.2.7). Wenn die Ebenen nicht bearbeitet worden sind, lässt sich das Originalbild wieder aus den Laplaceebenen herstellen.

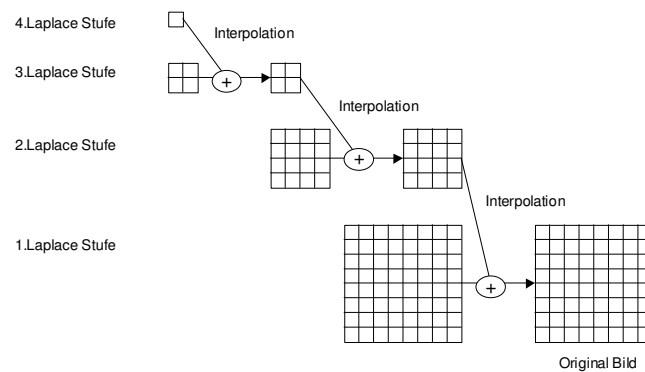


Abb. 1.2.11 : Schema der Rekonstruktion aus der Laplacepyramide

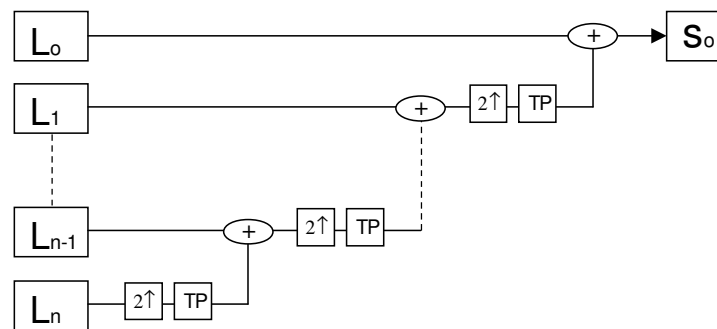


Abb. 1.2.12 : Flussdiagramm der Rekonstruktion

1.2.10 Binomialfilter

Das Filterdesign ist ein komplexes Optimierungsproblem. Die klassische Unschärferelation der Fouriertransformation verdeutlicht dies : Liegt im Ortsraum eine starke Lokalisierung vor, z.B. in Form einer Kante, so führt das im Frequenzraum (\rightarrow 1.2.13) zu einer Verschmierung der Funktion über einen weiten Bereich. Man muß also ein Filter finden, das in beiden Räumen ähnlich ist. Die Gaußsche Glockenfunktion ist neben dem Deltakamm die einzige Funktion, die in beiden Räumen die gleiche Form besitzt [4].

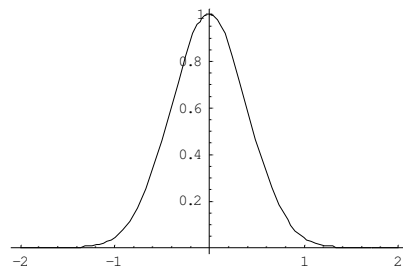


Abb. 1.2.13 : Zweidimensionale Gaussfunktion

Für positive Werte ist die Funktion monoton fallend. Sie fällt zuerst langsam, dann für höhere Werte schnell auf null ab.

Lässt man ein solches Filter über ein Bild laufen, so hat es eine glättende Wirkung (Tiefpass). Hohe Frequenzen werden herausgefiltert.

Auf einem diskreten Gitter approximiert man die Gaußsche Funktion durch die Binomialverteilung, die man mit dem Pascalschen Dreieck darstellen kann:

				1					0
			1		1				1
		1		2		1			2
	1		3		3		1		3
		1	4		6		4		4
	1		5		10		10		5
		1	6		15		20		6
	1		7		21		35		7
1		8		28		56		70	
	8		28		56		70		8
		28		56		70		56	
			28		56		70		
				28		56		70	
					28		56		
						28		56	
							28		
								28	
									28

Abb. 1.2.14 : Pascalsche Dreieck

1.2.11 Faltungsoperation

Bei einer diskreten Faltung von Bildpunkten werden Nachbarschaften eines Pixels miteinander verknüpft. Die Größe dieser Nachbarschaft wird in einem Fenster, oder auch Filtermaske, definiert. Die Form des Fensters und die Position des Pixels relativ zum Fenster muss angegeben werden. Werden die Bildpunkte im Bereich der Filtermaske mit einem entsprechenden Wichtungsfaktor der Maske multipliziert, die Produkte addiert und die Summe an die Stelle des entsprechenden Pixels geschrieben, spricht man von einer diskreten Faltung [5].

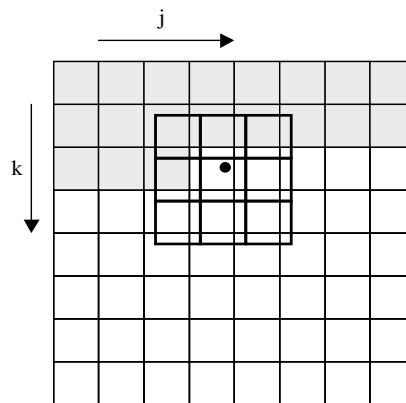


Abb. 1.2.15 : Prinzip einer Faltung, 3x3 Kern

In Abb.1.2.15 sind die grauunterlegten Pixel bereits durch die Faltungsoperation gelaufen. Daraus folgt, dass die bearbeiteten Bildpunkte in einem Zwischenspeicher abgelegt werden müssen. Die diskrete Faltung leitet sich aus der kontinuierlichen Faltung ab, die anstelle von Summenzeichen Integrale aufweist.

$$B(x, y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} b(\xi, \eta) h(x - \xi, y - \eta) d\xi d\eta$$

$$\Rightarrow B_{lm} = \sum_{k=0}^N \sum_{j=0}^M b_{jk} h_{l-j, m-k}$$

(1.2.9)

1.2.12 Separierbarkeit von Filtern

Der Vorteil der Separierbarkeit von Filtern liegt darin, dass man einen zweidimensionalen, linearen Filter auf zwei eindimensionale zurückführen kann. Hat man einen Filterkern mit k^2 Datenpunkten, so führt dies zu k^2 Multiplikationen und Additionen. Separiert man das Filter, so hat man nur jeweils k Multiplikationen und Additionen, also für 2 Raumrichtungen nur insgesamt $2k$ Operationen [8]:

Ein linearer Filter heißt separierbar, wenn für seine Impulsantwort gilt:

$h(x,y) = h_1(x) h_2(y)$, bzw. wenn für seine Übertragungsfunktion gilt

$$H(\omega_x, \omega_y) = H_1(\omega_x) H_2(\omega_y).$$

1.2.13 Impulsantwort und Übertragungsfunktion

Im Ortsraum wird ein linearer, diskreter Filter durch eine Faltung gegeben. Seine Eigenschaften sind vollständig durch die Koeffizienten gegeben. Die Filtermaske ist dabei identisch der Impulsantwort, d.h. die Antwort eines Filters auf ein Punktbild [5].

Ein Bild kann allerdings auch im Fourierraum, durch die s.g. Fouriertransformation, dargestellt werden. Auch hier ändert sich das Integralzeichen im Kontinuierlichen zu einem Summenzeichen im Diskreten:

$$F : \vec{w} \rightarrow \int_R \int_R f(\vec{x}) \cdot e^{-i \vec{w} \cdot \vec{x}} d\vec{x}$$
$$\Rightarrow A_{lm} = \sum_{j=0}^{N-1} e^{-i \frac{2\pi}{N} j l} \left[\sum_{k=0}^{M-1} a_{jk} e^{-i \frac{2\pi}{M} k m} \right] = \sum_{j=0}^{N-1} \sum_{k=0}^{M-1} a_{jk} e^{-i 2\pi \left[\frac{j l}{N} + \frac{k m}{M} \right]}$$

(1.2.10)

In diesem Raum führt sich die Faltung, laut dem Faltungstheorem, auf eine Multiplikation zurück. Die Fouriertransformierte der Faltungsmaske ist die sogenannte Übertragungsfunktion. Sie liefert für jede Wellenzahl einen Faktor, mit dem eine sich wiederholende Struktur durch die Filteroperation multipliziert wird. Bei diesem Faktor handelt es sich um eine komplexe Zahl, was dazu führt, dass eine periodische Struktur nicht nur eine Änderung in der Amplitude erfährt, sondern auch eine Phasenverschiebung.

1.3 Bisherige Methoden der Dynamikkompression

Besprochen werden an dieser Stelle nur Software-Methoden. Hardware-Lösungen finden in dieser Arbeit keine Erwähnung.

1.3.1 Zwei 8 bit Bilder

Wie bereits in der Einleitung erwähnt, wurde in der Vergangenheit das Problem der hohen Dynamik, d.h. der große Hub zwischen den relevanten Grauwertbereichen, dadurch gelöst, dass aus einem hochdynamischen Bild zwei 8 bit Bilder gemacht wurden, in denen jeweils nur die Fahrer, bzw. die Kennzeichen zu erkennen sind.



Abb. 1.3.1a : 8-bit Bild des dunklen Bereichs

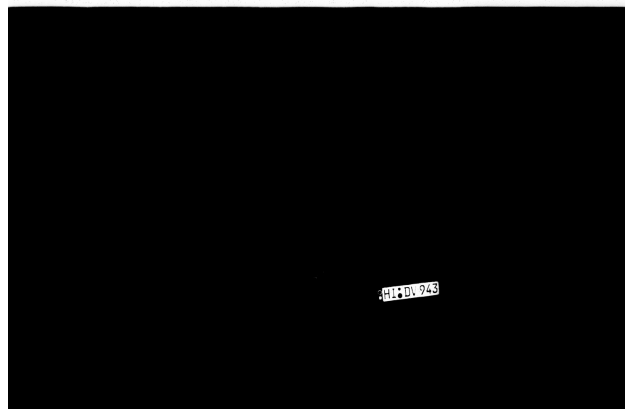


Abb. 1.3.1b : 8-bit Bild des hellen Bereichs

1.3.2 Gradationskurven in Photoshop

Photoshop bietet die Möglichkeit, mittels einer einstellbaren Gradationskurve, die Bereiche einer Komprimierung zu unterziehen, die dem Betrachter als unwichtig erscheinen, bzw. die Bereiche zu spreizen, die wichtige Details enthalten. Dies setzt natürlich voraus, dass die

Bereiche genau bekannt sind. Hier liegt auch der große Nachteil dieses Verfahren und macht es für die Dynamikkompression in der Verkehrsphotographie ungeeignet, da hier die Bereiche sehr schwanken. Vor allem bei den Reflexionen im Nummernschild spielen sehr viele Faktoren eine Rolle. Wird z.B. frontal auf das Nummernschild geblitzt, ergibt sich dort eine höhere Leuchtdichte, und damit ein höherer Grauwert, als wenn von der Seite geblitzt wird.

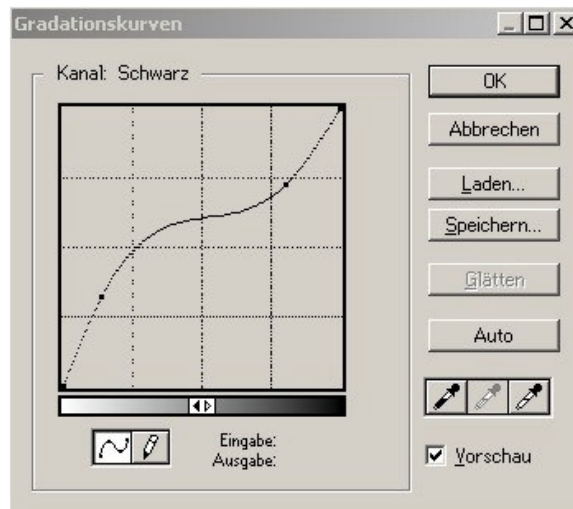


Abb. 1.3.2 : Gradationskurve in Photoshop

2 Material und Methode

2.1 Methode

Eine hohe Dynamik innerhalb der Verkehrsphotographie entsteht durch starke Reflexionen, die beim Blitzen vor allem im Bereich des Kennzeichens entstehen.

Die erste Überlegung bei der Dynamikkompression beschäftigt sich deshalb mit der reflektierenden Leuchtdichte, die einen multiplikativen Zusammenhang zwischen dem Remissionsgrad und der Beleuchtungsstärke aufweist. Durch eine Logarithmierung wird dieser Zusammenhang additiv.

Desweiteren wird, um die Helligkeit innerhalb des Nummernschildes zu reduzieren, eine Laplacepyramide berechnet, in der das Tiefpassbild mit einem Faktor kleiner eins belegt wird. Dieses Hochpassfilter hat die negative Eigenschaft, an steilen Kanten Überschwinger zu produzieren. In den Bandpassbildern sind vor allem die hohen Amplituden dafür verantwortlich, so dass sichergestellt werden muss, dass dieser Bereich linear behandelt wird.

Im Gegensatz dazu sollten kleine Amplituden verstärkt werden, weil sie hauptsächlich für lokale Kontraste zuständig sind; während ganze kleine Amplituden, die am meisten Rauschanteile besitzen, reduziert werden sollten.

Die Aufgabe des Parametertunings ist es, einen geeigneten Parametersatz für die nichtlineare Filterung innerhalb der Laplace Ebenen zu finden, so dass Überschwinger und Rauschen unterdrückt und lokale Kontraste verstärkt werden.



Abb. 2.1.1 : Blockbild des Dynamikkompressionsablaufs

2.2 Software

- 1.) WATCOM Integrated Development Enviroment Version 10.6;
C++ Compiler, mit dem die Programme pyramide.c und zeile.c programmiert wurden
- 2.) Bild.exe, Version 1.1 Copyright by Robot Visual Systems,
*Betrachtungsprogramm von Binärbildern, *.biu Bilder wurden mit diesem Programm als 16 bit Tiff Bilder gespeichert*
- 3.) Adobe Photoshop, Version 5.0
Mit Photoshop wurden die 16-bit Tiff Bilder in 8 bit Tiff Bilder umgewandelt und evtl. einer linearen Kontraststreckung unterzogen
- 4.) Zeile.exe
Hilfsprogramm, um die Pixelwerte einer beliebigen Zeile in ein Textfile auszugeben
- 5.) Microsoft Excel Tabellenkalkulation
Mit Excel wurden die Zeilendaten visuell in Diagramme umgesetzt
- 6.) Pyramide.exe
Eigenentwicklung im Rahmen der Diplomarbeit. Das Programm kann Bilder logarithmieren, in eine Laplacepyramide zerlegen und nach einer Filterung wieder ausgeben
- 7.) Adobe Pagemaker 6.5
DTP Programm, mit dem diese Arbeit gesetzt und erstellt wurde

2.3 Testbilder

Die von der Firma Robot Visual Systems zur Verfügung gestellten Bilder lagen alle im Binärformat vor (→ 5.5), d.h. die Daten sind weder komprimiert noch verschlüsselt. Dabei stehen jeweils 2 Bytes (16 bit) für ein Pixel. Am Anfang steht ein Header, der die X und Y-Dimension beinhaltet. Das Programm pyramide.c verarbeitet solche Bilder. Sie tragen das Kürzel *.biu (basic image unit).

Alle vorliegenden Bilder sind gescannte Analogbilder. Sie wurden mit einem Nikon 4000ED Scanner digitalisiert, sind schwarz/weiß und haben die Größe 3072 x 2048 Pixel, das entspricht 6.291.456 Pixel.

2.4 Darstellung der Bilder

Die Darstellung der Bilder dieser Arbeit setzt voraus, dass sie im 8 bit Format vorliegen, weil die meisten Viewer und DTP Programme nur 8 bit verarbeiten können. Deshalb ist es das Ziel, eine Konvertierung der Originalbilder auf 8 bit ohne Verlust relevanter Information zu erlangen. Zusätzlich setzt eine jpg. Komprimierung unter Photoshop eine 8 bit Form voraus.

Um dies zu erreichen, wurden die verarbeiteten Bilder im Programm Bild.exe im 16 bit Tiff-Format gespeichert und in Photoshop in ein 8 bit Tiff Bild umkonvertiert. Bei einigen Bildern wurde noch eine Kontraststreckung durchgeführt, um gewisse Teile besser hervorzuheben. Darauf wird an entsprechender Stelle hingewiesen.

Eine Alternative zu dieser Verarbeitung liegt darin, die Funktion „Auslesen“ in dem Programm pyramide.c durch die Funktion „Auslesen_8_bit“ zu ersetzen. Diese Funktion speichert das verarbeitete Bild direkt in einem 8 bit binären Format. Solche Bilder tragen das Kürzel *.bid. Das Problem bei dieser Methode liegt darin, dass bei fehlgeschlagener Verarbeitung, unter Umständen mit Informationsverlust zu rechnen ist.

Alle Bilder im Anhang sind mit Hilfe dieser Funktion ausgegeben.

Für Bilder aus der Gauss, bzw. aus der Laplacepyramide wurden im Programm pyramid.c zusätzliche Ausgabefunktionen implementiert.

2.5 Implementierung pyramide.c

2.5.1 Struktur des Programms

Das Programm pyramide.c ist das Kernstück der Verarbeitung. Es liest ein Bild im *.biu Format ein und logarithmiert es anschließend. Dann wird es in die einzelnen Laplaceebenen zerlegt in denen es nichtlinear gefiltert wird. Zum Schluss wird nach der Rekombination wieder ein *.biu File ausgegeben.

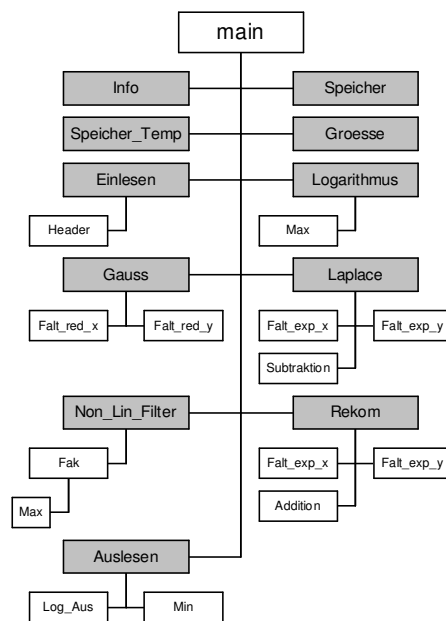


Abb. 2.5.1 : Struktur des Programms pyramide.c

2.5.2. Ein-/ Ausgabe der Parameter

Die Eingabe erfolgt über ein Textfile. Nach der Verarbeitung wird zusätzlich ein Logfile erstellt, in dem alle Parameter aufgelistet sind. Dies ermöglicht das Nachvollziehen der einzelnen Schritte beim Parametertuning.

In parameter.txt kann man folgende Parameter eingeben:

- Nummer des Bildes
- Namen des Bildes, welches bearbeitet werden soll
- die Entscheidung, ob eine Logarithmierung erfolgen soll (j/n)
- den linearen Bereich der Logarithmierung im oberen Bereich (prozentual zum Maximum)
- den Filterkern

- die Anzahl der Laplaceebenen
- den Skalierungsfaktor der jeweiligen Ebene
- den Exponenten der jeweiligen Ebene
- den nichtlinearen Bereich um den Nullpunkt der jeweiligen Ebene
- den Linearitätsbereich im oberen Bereich der jeweiligen Ebene
(prozentual zum Maximum)

Das Logfile beinhaltet zusätzlich noch den Zeitpunkt der Verarbeitung.

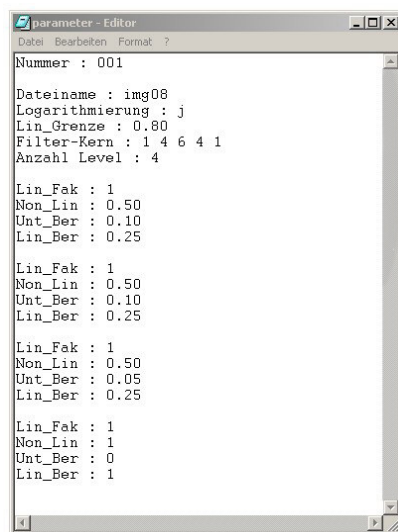


Abb. 2.5.2a : parameter.txt

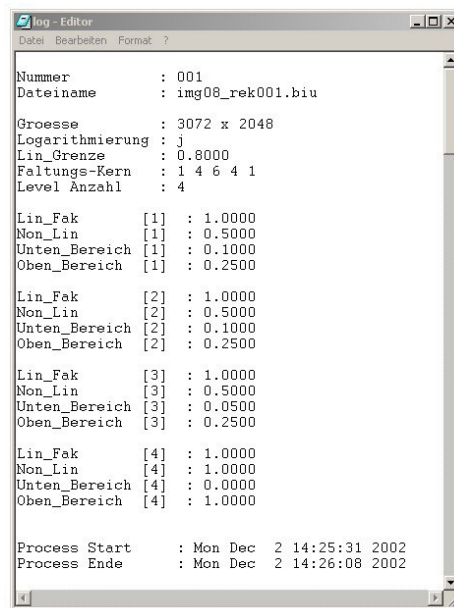


Abb. 2.5.2b : log.txt

2.5.3 Ablaufdiagramm

Für die Verarbeitung innerhalb des Programms wird eine pyramidenähnliche Speicherstruktur verwendet. Dabei ist die nächstkleinere Ebene immer um den Faktor 4 reduziert. Für die Zwischenspeicherung werden zusätzlich noch 2 Speicherplätze gebraucht.

Das Originalbild wird in die 0. Pyramidenebene geschrieben und dort auch logarithmiert. Die reduzierende Faltung, welche bei der Gausspyramide zum Einsatz kommt, benötigt nur die Speicherebene Temp[0] als Zwischenspeicher. Für die expandierende Faltung, die für die Laplaceberechnung und die Rekombination gebraucht wird, stehen zwei Zwischenspeicher zur Verfügung: Temp[0] und Temp[1]. Die Funktion „Fak“ greift auf alle Ebenen zu. Am Ende befindet sich das fertige Bild in der 0. Ebene der Pyramide, wo es auch ausgegeben wird.

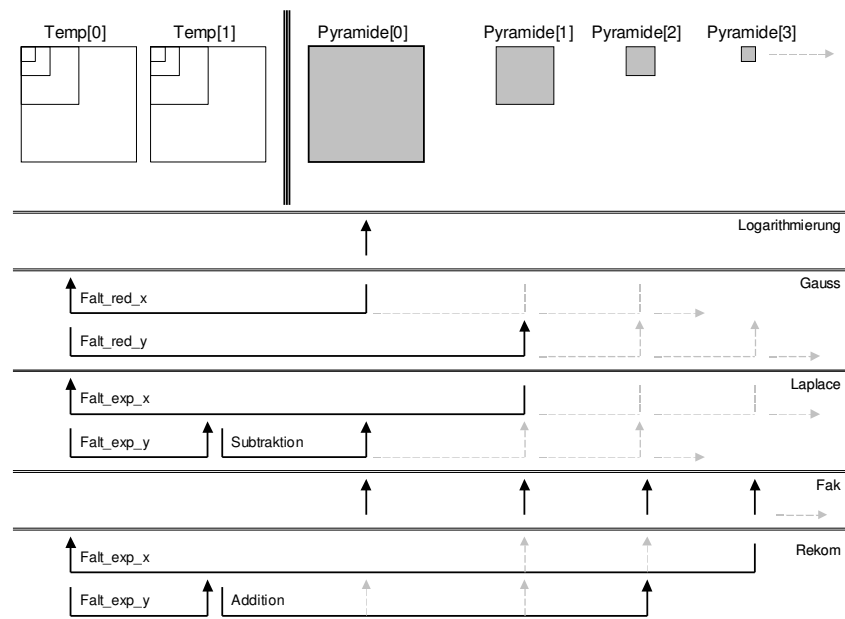


Abb. 2.5.3 : Ablaufdiagramm von `pyramide.c`

2.5.4 Reduzierende Faltung

Die reduzierende Faltung (\rightarrow 1.2.11) wird in zwei Schritten vollzogen, weil die Operation separierbar ist. Zuerst wird in x-Richtung gefaltet und dann in y-Richtung. Bei der Faltung in x-Richtung werden nur Pixel berechnet, die sich in einer geraden Spalte befinden, weil die Pixel in ungeraden Spalten später wegfallen, da das Bild bei der Gaußpyramide immer um den Faktor 2 in jeder Richtung unterabtastet wird. Deshalb braucht man bei der Faltung in y-Richtung auch nur die Pixel berücksichtigen, die sich in geraden Spalten und geraden Zeilen befinden. Für die Behandlung der Ränder wird eine Spiegelung implementiert. Ein 5x5 Kern braucht jeweils 2 zusätzliche Spalten, bzw. Zeilen. Dabei wird bei der Faltung in x-Richtung die 3. Spalte in die 1. Spalte und die 4. Spalte in die 0. Spalte, bzw. die (x_dim) . Spalte in die (x_dim+2) . Spalte und die (x_dim-1) . Spalte in die (x_dim+3) . Spalte kopiert. Bei der Faltung in y-Richtung wird genauso verfahren, wobei hier zusätzlich auf die Spiegelung der (y_dim-1) . Zeile verzichtet werden kann.

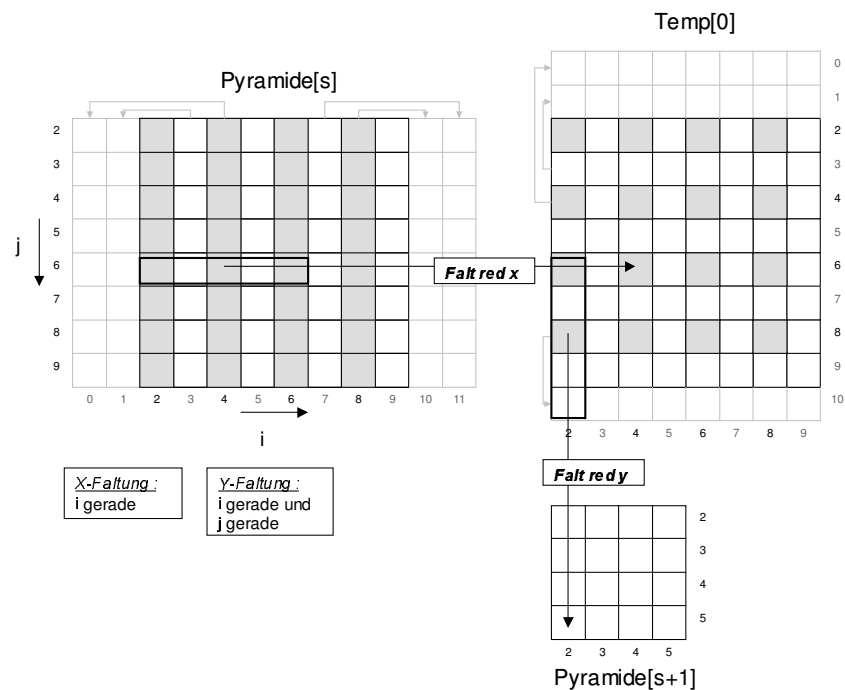


Abb. 2.5.4 : Ablauf der reduzierenden Faltung

2.5.5 Expandierende Faltung

Bei der expandierenden Faltung (\rightarrow 1.2.11) wird auch wie bei der reduzierenden Faltung jede Richtung separat durchgeführt. Man muss verschiedene Pixel unterscheiden, da man sich das Bild, bevor es gefaltet wird, mit Nullen aufgefüllt vorstellen muss. Bei der Faltung in x-Richtung brauchen Pixel mit gerader Spalten- und Zeilenzahl nur 3 relevante Elemente, während Pixel mit ungerader Spaltenzahl aber gerader Zeilenzahl nur 2 relevante Elemente benötigen. Pixel mit ungerader Zeilenzahl werden Null gesetzt. Bei der Faltung in y-Richtung benötigen Pixel mit gerader Zeilenzahl 3 relevante Elemente und Pixel mit ungerader Zeilenzahl 2 relevante Elemente. Die Spiegelung ist analog zur Spiegelung der reduzierenden Faltung.

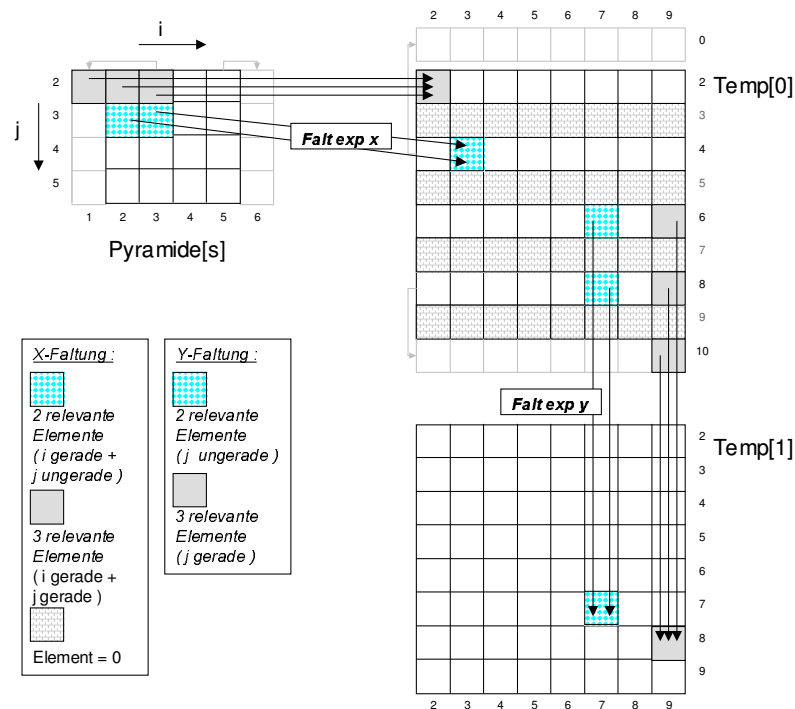


Abb. 2.5.5 : Ablauf der expandierenden Faltung

2.5.6 Logarithmierung

Bei der Logarithmierung kann man in der Datei parameter.txt den Wert Lin_Grenze eingeben. Er bestimmt die Stelle, ab der die Werte nicht mehr logarithmiert werden, sondern linear auf sich selbst abgebildet werden. Dies ist nötig, um im Kennzeichen einen hohen Kontrast zu bewahren (\rightarrow 3.1). Der Wert bezieht sich dabei prozentual auf das Maximum im Bild. Im Argument des Logarithmus steht $x+1$ damit kein $\log(0)$ auftritt.

$$f(x) = c \cdot \log(x + 1) \quad \text{für } x < x_g$$

$$f(x) = x \quad \text{für } x \geq x_g$$

(2.5.1)

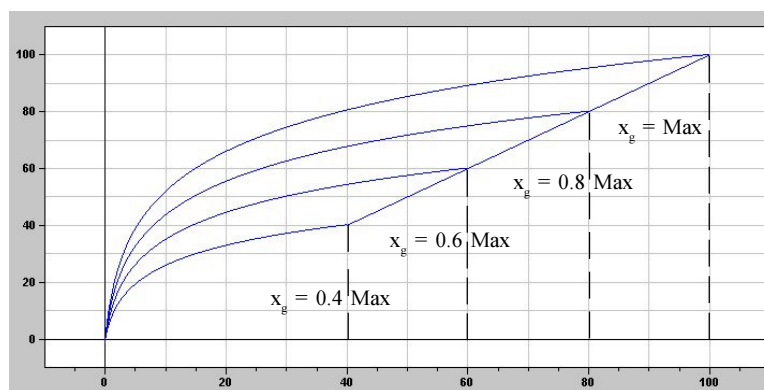


Abb. 2.5.6 : Beispiele für den Parameter Lin_Grenze

2.5.7 Filterung in den Laplaceebenen

Bei der eigentlichen Filterung in den Ebenen der Laplacepyramide wird mit folgender zusammengesetzter Funktion (2.5.2) gearbeitet.

$$f(x) = a \cdot \frac{x}{|x|} \cdot x_c \cdot \left(\frac{x}{x_c} \right)^{\frac{1}{p}} \quad \text{für } |x| \leq x_c$$

$$f(x) = a \cdot \frac{x}{|x|} \cdot (M - x_c) \cdot \left(\frac{|x| - x_c}{M - x_c} \right)^p + x_c \quad \text{für } x_c < x < M$$

$$f(x) = a \cdot \frac{x}{|x|} \cdot (M - x_c) \cdot \left(\frac{|x| - x_c}{M - x_c} \right)^p - x_c \quad \text{für } -M < x < -x_c$$

$$f(x) = a \cdot x \quad \text{für } |x| \geq M$$

$$\text{mit } x_c \leq M \leq \text{Max} \quad (2.5.2)$$

Um den Nullpunkt liegt die Funktion unterhalb einer Geraden mit der Steigung 1, d.h. es findet keine Verstärkung statt. Hier sollten sich die ganzen kleinen Amplituden des Rauschens befinden. Nach diesem Bereich liegt die Funktion oberhalb einer Geraden mit der Steigung 1. Hier findet eine Verstärkung statt. Anschließend geht die Funktion in eine Gerade über.

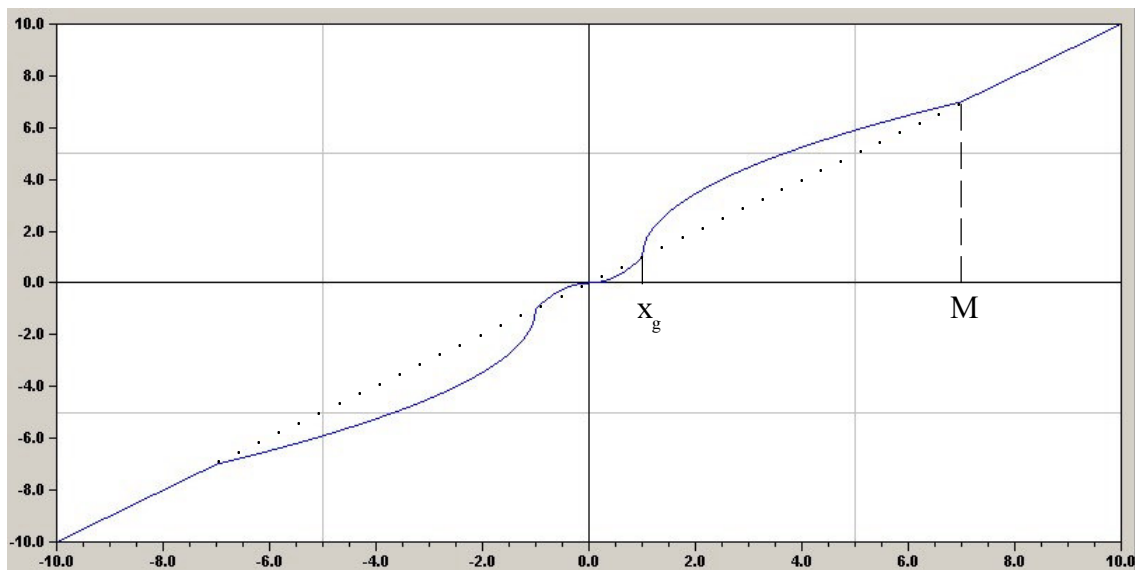


Abb. 2.5.7 : Filterfunktion innerhalb der Laplaceebenen

Formel	Bezeichnung	Bezeichnung in parameter.txt
a	Skalierungsfaktor	Lin_Fak
p	Exponent	Non_Lin
x _c	unterer nichtlinearer Bereich	Unt_Ber
M	obere Linearitätsgrenze	Lin_Ber
Max	absolutes Maximum	

M bezieht sich dabei prozentual auf Max und x_c prozentual auf M.

Abb. 2.5.8. : Tabelle über einstellbare Parameter

2.5.8 Filterkern

In der Datei parameter.txt (→ 2.5.2) kann man die Koeffizienten eines 5x5 Filterkerns für die Faltung zur Berechnung der Gauss-, bzw. der Laplacepyramide frei eingeben.

Hieraus berechnet sich im Programm ein Ausgleichsfaktor, mit dem das Bild bei der Expansion multipliziert werden muss, nachdem es mit Nullen aufgefüllt wird, um den Helligkeitsbereich des Originalbildes zu erhalten. Bei einem Binomialfilter (→ 1.2.10) lautet der Faktor, wie bereits erwähnt, 2 pro Raumrichtung. Oder anders ausgedrückt muß gelten:

$$H(0) = 2 \text{ und } H(\pi) = 0$$

π stellt dabei die Nyquistfrequenz dar.

Ein Rechteckfilter erfüllt diese Bedingung nicht. Geht man von der Faltung in X-Richtung aus, hat man Pixel, für die 3 Pixel relevant sind. Das sind solche, die an einer geraden X- und Y-Position stehen. Hier werden nur die Koeffizienten 1, 3 und 5 benutzt, weil die anderen durch das Auffüllen mit Nullen wegfallen.

Hat man nun einen 1 1 1 1 1 Kern, so ist die gesamte Summe 5. Da nur der 1., 3. und 5. Koeffizient zum Tragen kommt, errechnet sich daraus ein Ausgleichfaktor von 5/3. Analog ergibt sich für Pixel an ungerader X-Position und gerader Y-Position einen Faktor von 5/2.

Daraus ergibt sich dann ein Kern der Form $\frac{1}{3} \begin{pmatrix} 1 & 1 & 1 \\ 2 & 3 & 2 \end{pmatrix} \frac{1}{3}$.

Zusätzlich muß gelten, dass der Kern normiert und symmetrisch ist.

In weiteren Verlauf des Parametertunings wird deshalb ein Binomialfilter der Größe 5 benutzt :

$$h = \frac{1}{16} \cdot (1 \ 4 \ 6 \ 4 \ 1)$$

3 Testergebnisse

3.1 Logarithmierung

Die Logarithmierung wird eingesetzt, um das Produkt „Beleuchtungsstärke x Reflexion“ in einen additiven Zusammenhang umzuwandeln (\rightarrow 1.2.4). Benutzt man nur eine Abbildung in der Form $\log(x+1)$, kann es sehr schnell passieren, dass die Informationen des überstrahlten Nummernschildes im logarithmierten Bild einen zu kleinen Grauwertbereich abdecken, so dass ein hoher Kontrast nach der Verarbeitung nicht mehr gewährleistet ist. Deshalb wird ein Teil des Grauwertbereichs auf sich selbst abgebildet. Dazu kann man in den Parametern die Größe Lin_Grenze angeben, die einen prozentualen Anteil des absoluten Maximums angibt, ab dem die Werte keine Logarithmierung mehr erfahren.

Die Bereiche, in denen sich die Informationen des Kennzeichen befinden, erstrecken sich im Originalbild durchschnittlich über einen Bereich von ca. 13.000 Graustufen. Da sich diese Grauwerte im hohen Helligkeitsbereich, aufgrund der Überstrahlung, befinden, bildet eine Logarithmierung diesen Bereich auf ca. 5.000 Grauwerte ab. Dieser Bereich reicht nicht aus, um einen hohen Kontrast nach der Verarbeitung in einer Laplacepyramide sicherzustellen. Wird jedoch ein Bereich definiert, in dem die Pixel linear verarbeitet werden, steigt dieser Bereich wieder an. Bei verschiedenen Tests hat sich ein Wert zwischen 0.75 und 0.80 bewährt, d.h. 20 bis 25% unterhalb des absoluten Maximums werden nicht logarithmiert. Mit einem Wert von 0.80 ergibt sich ein Bereich von ca. 11.000 Graustufen für die Informationen des Kennzeichens, so dass ein hoher Kontrast nach der Filterung in den Laplacepyramiden sicherstellt ist.



Abb. 3.1.1: Logarithmiertes Bild in Bild.exe als 16 bit Tiff gespeichert; anschließend in Photoshop einer Autotonwert-Korrektur unterzogen und als 8 bit Tiff abgespeichert.



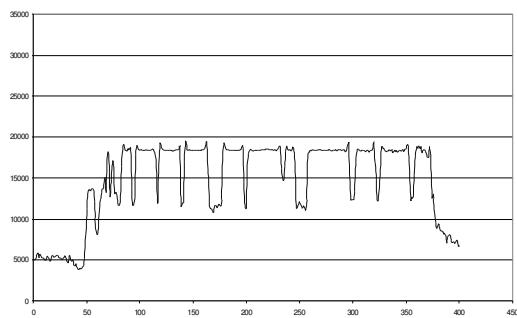
Abb. 3.1.2 : wie Abb.3.1.1 nur mit Lin_Grenze = 0.80

Man erkennt in den Abb. 3.1.1 und 3.1.2, dass der Kontrast im Nummerschild mit zunehmenden linearen Bereich steigt. Dies hat zur Folge, dass der Grauwertbereich, in dem der Fahrer zu erkennen ist, stärker komprimiert wird. Somit hat die Größe Lin_Grenze einen direkten Einfluss darauf, welches Element (Kennzeichen oder Fahrer) letztendlich einen höheren Kontrast aufweist.

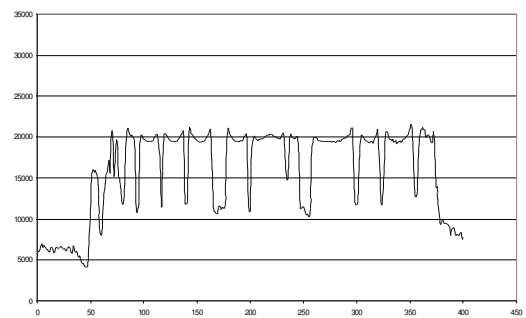
Die Logarithmierung ist das erste Element, welches zur Dynamikkompression beiträgt, weil es Bereiche mit niedrigen Grauwerten spreizt und höhere Bereiche komprimiert.

3.2 Wahl der Anzahl der Laplacestufen

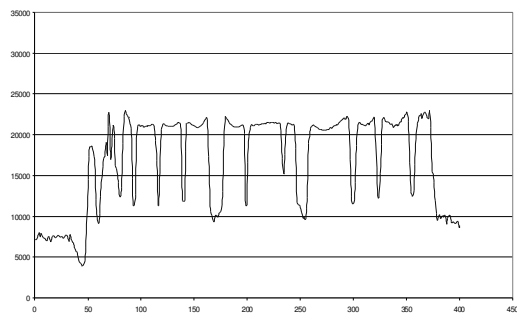
Das zweite Mittel, um die Dynamik zu komprimieren, beruht darauf, dass man das Tiefpassbild in der Laplacepyramide mit einem Faktor kleiner 1 belegt, und so die niederfrequenten Anteile des Bildes reduziert. Die erste Frage, die sich in diesem Zusammenhang stellt, ist die Überlegung, wieviele Stufen berechnet werden müssen, um das richtige Tiefpassbild zu erhalten. Dazu wurde eine Testreihe angefertigt, in der jeweils die Anzahl der Stufen um 1 erhöht wurde. Das Tiefpassbild erfuhr anschließend eine Multiplikation mit dem Faktor 0.5. Die Bandpassbilder unterlagen dabei keiner Filterung.



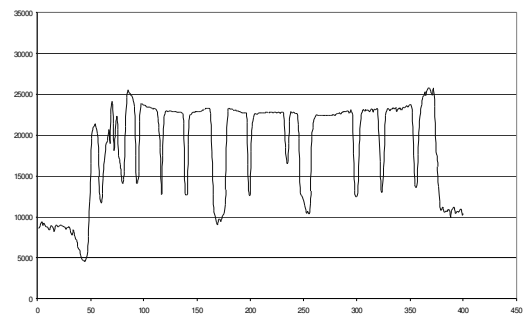
2 Ebenen



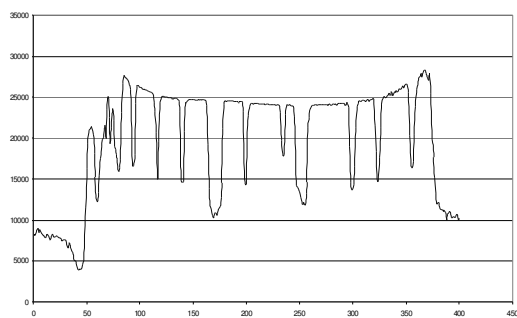
3 Ebenen



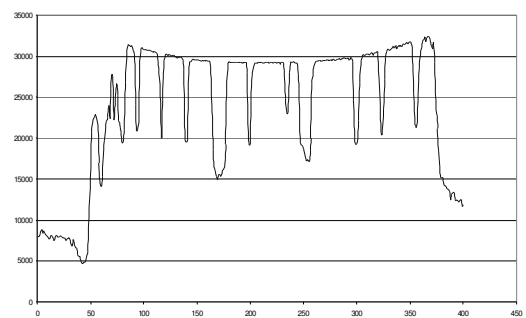
4 Ebenen



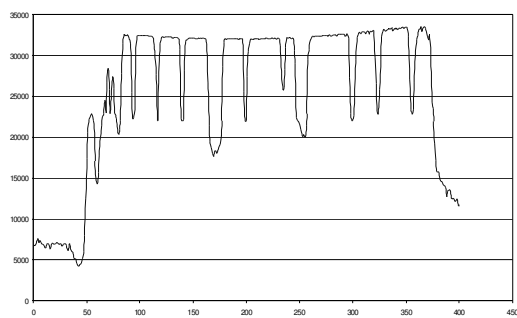
5 Ebenen



6 Ebenen



7 Ebenen



8 Ebenen

Abb. 3.2.1 : Vergleich eines Schnittes durch ein Kennzeichen in Abhängigkeit von der Anzahl der Laplaceebenen

Wie man erkennt, bilden sich lokale Überschwinger, wenn man nur mit max. 4 Ebenen arbeitet. Bis zur 7. Ebene tritt ein anderes Phänomen auf: Dort hat man einen Helligkeitsverlust innerhalb des Kennzeichens. Ab der 8. Ebene reguliert sich dieser Verlust wieder.

Da das Tiefpassbild der Laplacepyramide immer das letzte Bild der Gausspyramide ist, kann man aus der Gausspyramide erkennen, wie das niederfrequente Bild in Abhängigkeit von der Anzahl der Ebenen aussieht. So erfährt man, welche Inhalte man reduziert, indem man das Tiefpassbild mit einem Faktor kleiner 1 belegt.



Abb. 3.2.2a : Tiefpass, 2 Ebenen



Abb. 3.2.2b : Tiefpass, 3 Ebenen



Abb. 3.2.2c : Tiefpass, 4 Ebenen



Abb. 3.2.2d : Tiefpass, 5 Ebenen



Abb. 3.2.2e : Tiefpass, 6 Ebenen

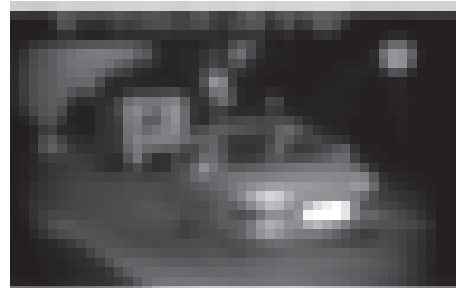


Abb. 3.2.2f : Tiefpass, 7 Ebenen

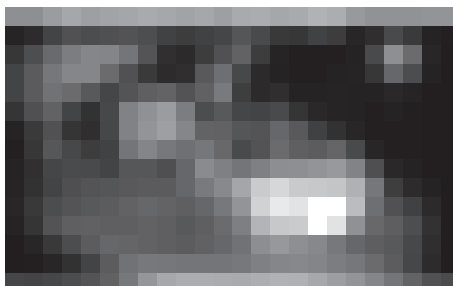


Abb. 3.2.2g : Tiefpass, 8 Ebenen

Um sicherzustellen, Überschwinger und Helligkeitsschwankungen, speziell innerhalb des Kennzeichens, so gering wie möglich zu halten, und auch bei der Reduktion der tieffrequenten Anteile nur die Helligkeit im Nummernschild zu reduzieren, werden für den weiteren Prozess jeweils 8 Ebenen berechnet.

3.3 Lineare Filterung

Um die Dynamik zu komprimieren, d.h. um alle relevanten Bildelemente, wie Fahrer und Kennzeichen, in einen Bereich von ca. 8 bit zu transformieren, wird das Tiefpassbild mit einem Faktor kleiner 1 reduziert. Zunächst wird nur linear gefiltert, ohne auf die Problematik von Überschwängern oder Rauschen (\rightarrow 1.2.5) einzugehen. Ausgegangen wird von einem typischen Bild, in dem sich die wichtigen Informationen in einem Bereich befinden, der sich weit über 8 Bit erstreckt:



Abb. 3.3.1 : Originalbild, dunkler Bereich (unteren 8-bit)

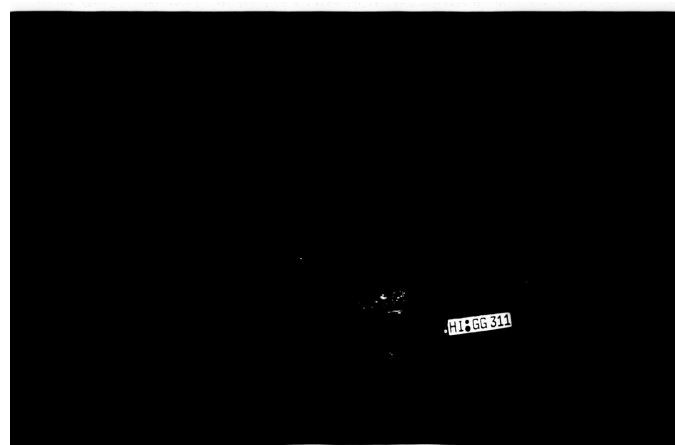


Abb. 3.3.2 : Originalbild, heller Bereich (oberen 8-bit)



Abb. 3.3.3 : Linear gefiltertes Bild, Multiplikation des Tiefpass Bildes mit 0.5



Abb. 3.3.4 : Linear gefiltertes Bild, Multiplikation des Tiefpass Bildes mit 0.8

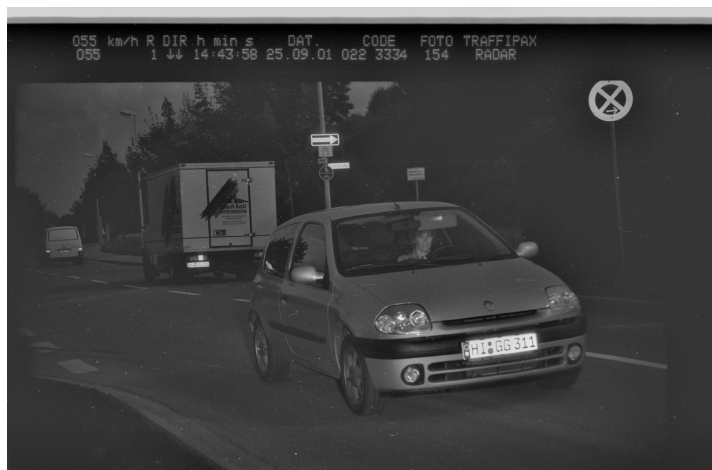


Abb. 3.3.5 : Linear gefiltertes Bild, Multiplikation des Tiefpass Bildes mit 0.3

Wie man erkennt, nimmt der globale Kontrast (und die globale Helligkeit) in Abhängigkeit von der Stärke der Reduktion des Tiefpassbildes ab. Hochauflösende lokale Kontraste können mit diesem Vorgehen nicht beeinflusst werden.

3.4 Nichtlineare Filterung

Mit Hilfe einer nichtlinearen Filterung (\rightarrow 2.5.7) innerhalb der Laplaceebenen kann man erreichen, dass Überschwinger und Rauschen reduziert werden und vor allem lokale Kontraste, wie z.B. im Gesicht erhöht werden. Dem Ganzen liegt folgende Idee zu Grunde:

Überschwinger, die in der Regel eine hohe Amplitude aufweisen, werden genauso wie die ganz kleinen Amplituden des Rauschens nicht verstärkt, bzw. abgeschwächt, während kleine Amplituden, die wichtige Details aufweisen, verstärkt werden. Dazu ist es sinnvoll, sich einmal einen Schnitt durch eine Laplaceebene anzugucken. Da vor allem das Kennzeichen und das Gesicht des Fahrers später die beste Qualitäten aufweisen sollen, ist es sinnvoll, auch diese Bereiche für genauere Untersuchungen herauszunehmen:

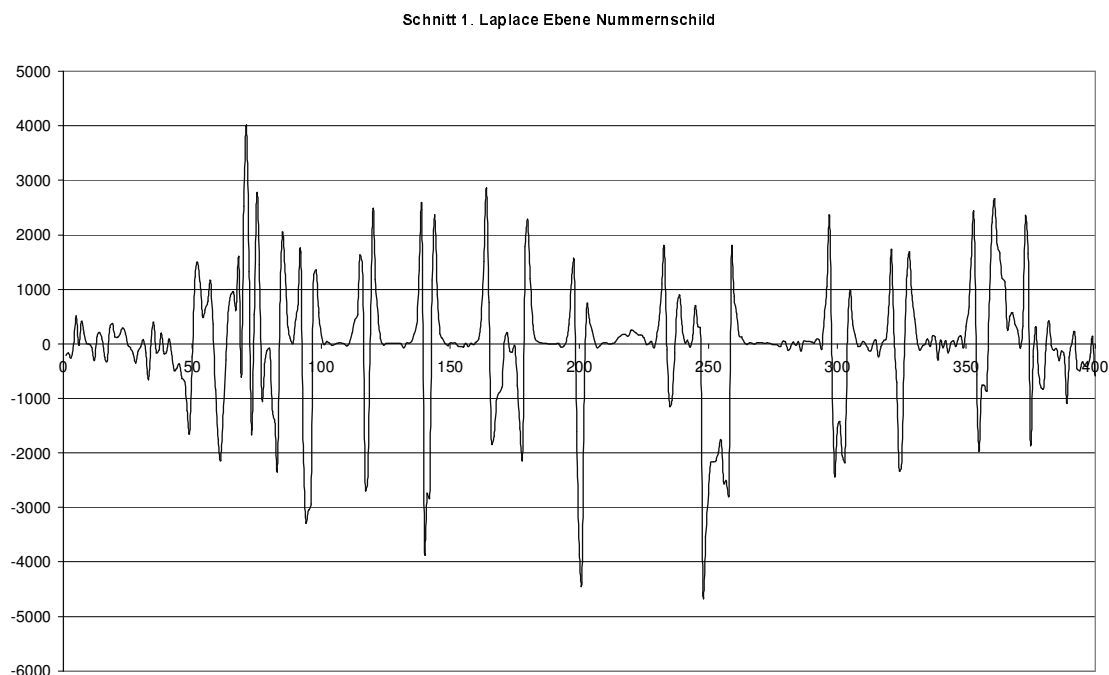


Abb. 3.4.1 : Schnitt durch ein Nummernschild im Hochpassbild der Laplaceebene

Was man in Abb.3.4.1 erkennt, sind Überschwinger an den Helligkeitssprüngen, die die Buchstaben auf dem Nummernschild zu dem Hintergrund bilden. Diese Bereiche sollten ebenso wie die ganz kleinen Amplituden innerhalb der Buchstaben, die einen gewissen Rauschpegel darstellen, nicht verstärkt werden.

Dazwischen sollte eine Verstärkung der Amplituden realisiert werden, die für den Kontrast innerhalb des Fahrers zuständig sind.

Die Aufgabe des Parametertunings liegt nun darin, die vier zur Verfügung stehenden Größen optimal an das Bildmaterial anzupassen.

- Mit der Grenze des unteren nichtlinearen Bereichs muss sichergestellt werden, dass das Rauschen nicht mitverstärkt wird.
- Bei der Grenze des oberen linearen Bereichs muss darauf geachtet werden, dass die Überschwinger nicht verstärkt werden.
- Mit der Wahl des Exponenten stellt man die Grenze und den Grad der Verstärkung innerhalb des nichtlinearen Teils ein, in dem sich die Bereiche, wie z.B. die des Gesichts befinden.
- Mit dem Skalierungsfaktor muss man die Tiefpassebene geeignet mit einem Faktor belegen, um die gewünschte Dynamikkompression, bzw. den gewünschten globalen Kontrast zu erhalten.

Zur Veranschaulichung des Problems wird der Bereich des Gesichts ein wenig genauer betrachtet:

Das lineare Behandeln der einzelnen Laplaceebenen (\rightarrow 3.3) hat bereits in der Dynamikfrage einige gute Ergebnisse erzielt. Was bei diesem Vorgehen jedoch nicht behandelt werden konnte, war die Detailverbesserung innerhalb des Fahrers.



Abb.3.4.2 : Lineare Filterung, Faktor der Tiefpassebene = 0.5

Bei den folgenden Bildern kamen 8 Laplaceebenen zum Einsatz. Der lineare Bereich für die Logarithmierung wurde mit 0.80, der Exponent mit 0.5 und der Faktor des Tiefpassbildes mit 0.5 gewählt. Der obere lineare Bereich wurde auf 0.75 des eigentlichen Maximums festgelegt. Variiert wurden nur die unteren nichtlinearen Bereiche der ersten beiden Laplaceebenen. Für die restlichen Ebenen wurde ein Wert von 0.05 benutzt.



Abb. 3.4.3a



Abb. 3.4.3b



Abb. 3.4.3c

Abb. 3.4.3 : Details innerhalb des Gesichtes

a: Unterer Nichtlinearitätsbereich der 1. Ebene : 0.05 und der 2. Ebene : 0.05

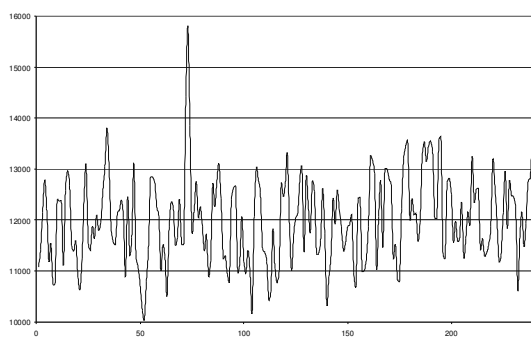
b: Unterer Nichtlinearitätsbereich der 1. Ebene : 0.10 und der 2. Ebene : 0.10

c: Unterer Nichtlinearitätsbereich der 1. Ebene : 0.20 und der 2. Ebene : 0.20

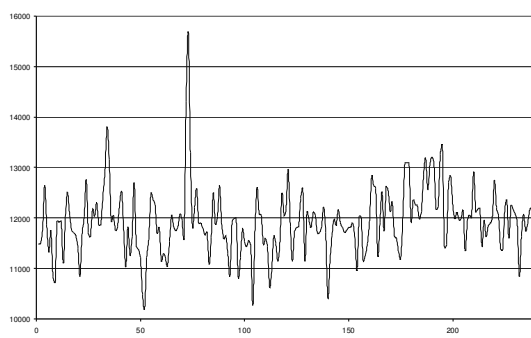
In Abb. 3.4.3 ist zu erkennen, dass der Kontrast innerhalb feiner Strukturen deutlich gegenüber der linearen Filterung zunimmt. Als Nebeneffekt tritt hier eine Verstärkung des ebenfalls hochfrequenten Rauschens auf, wenn der untere nichtlineare Bereich nicht richtig gewählt wird. Ist er nicht ideal bestimmt, kann es wie in Abb. 3.4.3c auch zu Unschärfe kommen. Der Effekt des Rauschens wird vor allem auf der Motorhaube sehr deutlich, da es sich hier um eine relativ homogene Fläche handelt. Um dies ein wenig anschaulich zu machen, wurde jeweils ein Schnitt durch die Bilder gelegt, die auch schon in Abb.3.4.3 verwendet wurden:



Abb.3.4.4 : Markierung des Schnitts durch die Motorhaube



*Abb. 3.4.5: Schnitt durch die Motorhaube
Parameter wie in Abb. 3.4.3a*



*Abb. 3.4.6 : Schnitt durch die Motorhaube
Parameter wie Abb. in 3.4.3b*

In den Abbildungen 3.4.5 und 3.4.6 erkennt man, dass die Höhen der kleinen Amplituden, und damit das Rauschen, von dem Bereich des unteren definierten Komprimierungsbereichs abhängen. Je größer dieser Bereich gewählt wird, desto weniger hochfrequente Anteile werden verstärkt. Wird der Bereich zu groß gewählt, kann es passieren, dass Strukturen verlaufen (→ Abb. 3.4.3c).

Um einen guten Bildeindruck zu erreichen, dürfen die Überschwinger in dem Kantenbild, also dem Hochpassbild, nicht mit verstärkt werden. Dazu muß der Parameter `Oben_Lin` ideal gewählt werden.



Abb.3.4.7 : Überschwinger im Kantenbild der Laplaceebene

Wählt man diesen Wert passend aus; bei den getesteten Bilder bewährte sich ein Wert von 0.25, d.h. 75% des absoluten Maximums werden linear behandelt, so kann man im rekombinierten Bild kaum Überschwinger erkennen:

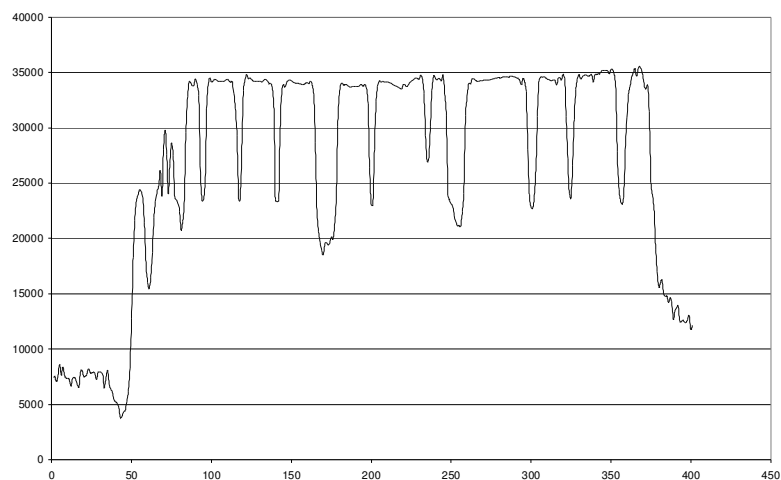


Abb. 3.4.8 : Helligkeitsverlauf innerhalb des Nummernschildes nach optimaler Filterung

In den Abb. 3.4.9 und 3.4.10 sieht man eine Skriptzeile, die immer mit aufs Bild belichtet wird, und alle Informationen der Aufnahmesituation enthält. Man kann feststellen, dass sie durch die Filterung in den Laplaceebenen weder im Kontrast noch in der Lesbarkeit verändert wird. Sie wird genau wie alle wichtigen Elemente in eine 8-bit Umgebung hineintransformiert, ohne Informationsverlust zu erleiden.



Abb.3.4.9 : Skriptzeile im Originalbild (Darstellung der Graustufen 0 - 13.000)

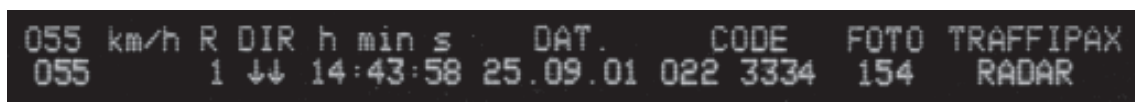


Abb.3.4.10 : Skriptzeile im rekombiniertem 8-bit Bild

Wählt man einen idealen Parametersatz, so werden wie in der Abb. 3.4.11 zu sehen ist, die ganz kleinen Amplituden abgeschwächt, die kleinen Amplituden verstärkt und die großen Amplituden beibehalten.

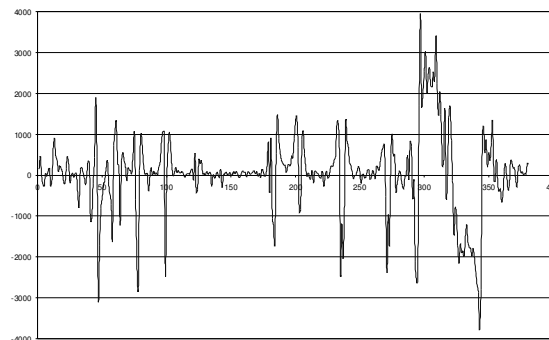


Abb. 3.4.11a : Querschnitt Kennzeichen in der 1. Laplaceebene vor der Filterung

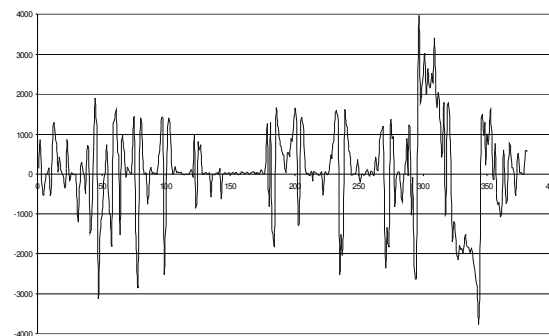


Abb. 3.4.11b : Querschnitt Kennzeichen in der 1. Laplaceebene nach der Filterung

Zusammenfassend kann man sagen, dass sich folgender Parametersatz, für die vorliegenden Testbilder bestens eignet, gute Ergebnisse zu erzielen:

Logarithmierung	:	ja
Lin_Grenze	:	0.8
Faltungskern	:	1 4 6 4 1
Level Anzahl	:	8
Lin_Fak	[1] :	1.00
Non_Lin	[1] :	0.50
Unt_Ber	[1] :	0.10
Lin_Ber	[1] :	0.25
Lin_Fak	[2] :	1.00
Non_Lin	[2] :	0.50
Unt_Ber	[2] :	0.10
Lin_Ber	[2] :	0.25
Lin_Fak	[3] :	1.00
Non_Lin	[3] :	0.50
Unt_Ber	[3] :	0.05
Lin_Ber	[3] :	0.25
Lin_Fak	[4] :	1.00
Non_Lin	[4] :	0.50
Unt_Ber	[4] :	0.05
Lin_Ber	[4] :	0.25
Lin_Fak	[5] :	1.00
Non_Lin	[5] :	0.50
Unt_Ber	[5] :	0.05
Lin_Ber	[5] :	0.25
Lin_Fak	[6] :	1.00
Non_Lin	[6] :	0.50
Unt_Ber	[6] :	0.05
Lin_Ber	[6] :	0.25
Lin_Fak	[7] :	1.00
Non_Lin	[7] :	0.50
Unten_Ber	[7] :	0.05
Lin_Ber	[7] :	0.25
Lin_Fak	[8] :	0.50
Non_Lin	[8] :	1.00
Unt_Ber	[8] :	0.00
Lin_Ber	[8] :	1.00

Abb. 3.4.12 : Tabelle mit idealem Parametersatz

4 Diskussion

Das Verfahren der nichtlinearen Filterung in einer Laplacepyramide mit vorangehender Logarithmierung ist ein sehr gutes Verfahren, um die Dynamik im Rahmen der Verkehrsphotographie zu komprimieren. Zusätzlich kann mit dem geeigneten Parametersatz sehr gut Einfluss auf Überschwinger und das Rauschverhalten genommen werden. Während die Tiefpassfilterung in der Laplacepyramide und die Logarithmierung relativ unabhängig von den vorliegenden Belichtungsverhältnissen ist, müssen die Parameter, die zur Unterdrückung von Überschwüngen und Rauschen eingestellt werden, doch wesentlich näher an die vorhandenen Gegebenheiten angepasst werden. Dennoch ist das ganze Verfahren soweit unabhängig, als dass es für alle zur Verfügung stehenden Bilder mit dem gleichen Parametersatz aus 3.4 sehr gute Ergebnisse liefert:



Abb 4.1 : Originalbild, reguläre Belichtung, Darstellung der dunklen 8 bit

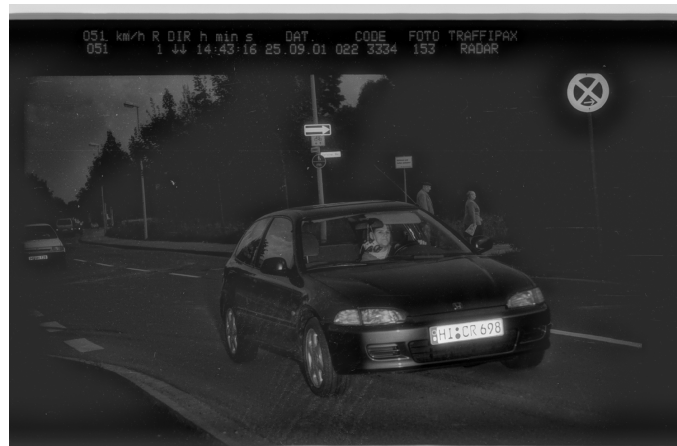


Abb. 4.2 : Gefiltertes 8 bit Bild nach dem Parametersatz aus 3.4

(die Artefakte im unteren mittleren Bereich sind bereits auf dem Negativ vorhanden)



Abb. 4.3 : Originalbild, Reflexionen des Nummernschildes gehen an Kamera vorbei



*Abb. 4.4 : Gefiltertes 8-bit Bild nach dem Parametersatz aus 3.4
(die kreisrunden Artefakte sind breits auf dem Negativ vorhanden)*

Wie in den Abb. 4.1 bis 4.4 zu erkennen ist, funktioniert der gleiche Parametersatz auch für unterschiedlich belichtete Bilder. Beim ersten Bild ist das Nummernschild vollkommen überstrahlt, während im zweiten Bild die Reflexionen an der Kamera vorbei gehen, und so nicht für einen großen Hub in der Helligkeit zwischen Fahrer und Kennzeichen sorgen. In beiden Fällen wurden alle wichtigen Elemente in einen 8 bit Bereich transformiert. Selbst der subjektive Eindruck der Bilder konnte in beiden Fällen verbessert werden.

Problematisch wird es nur, wenn die Aufnahme fehlschlägt und sich dadurch die Informationen des Nummernschildes nur in einem sehr schmalen Grauwertbereich befinden. In diesem Fall kann es passieren, dass diese Informationen zwar nicht verloren gehen, sie aber nicht mehr zusammen mit dem Fahrer in einem Bild zusammen sichtbar gemacht werden können, weil der Kontrast innerhalb des Nummernschildes zu gering ist.

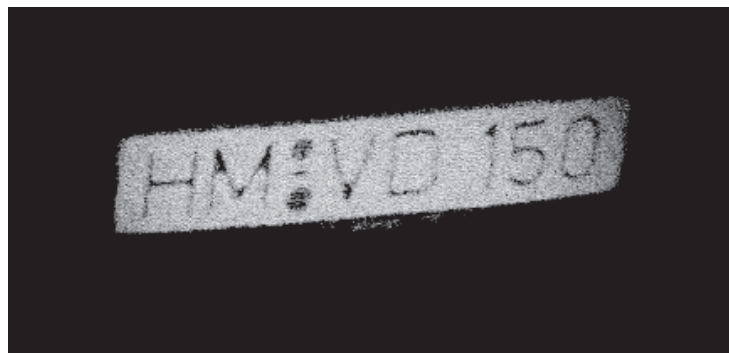


Abb. 4.5 : Kennzeichen im Originalbild, Darstellung der Grauwerte 63.000 bis 65.000



*Abb. 4.6 : Gefiltertes Bild mit dem Parametersatz aus 3.4;
in der Darstellung des unteren Kennzeichen wurde der Kontrast durch eine lineare
Spreizung verbessert*

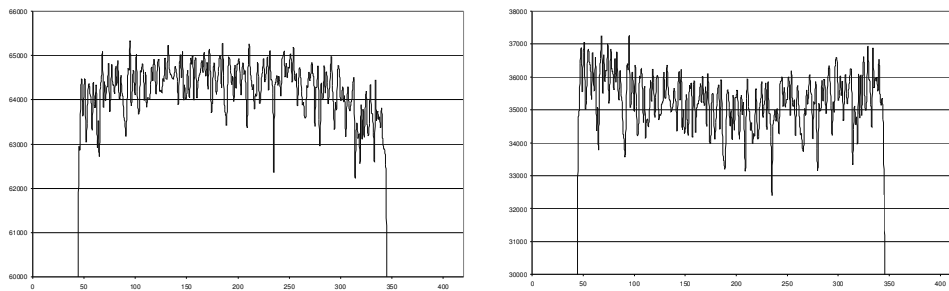


Abb.4.7 : Schnitt durch das Kennzeichen im Originalbild (links) und im rekombinierten Bild (rechts); (man erkennt, der geringe, rauschende Kontrast bleibt erhalten)

4.1 Zusammenfassung

Zusammenfassend kann man sagen, dass das Verfahren der nichtlinearen Filterung innerhalb einer Laplacepyramide mit vorangehender Logarithmierung, unter allen getesteten Belichtungsverhältnissen sehr gute Ergebnisse liefert. Aufgrund dessen, dass immer die gleichen Parameter eingesetzt werden können, ist das Programm für eine automatische Verarbeitung geeignet.

Desweiteren läuft man keine Gefahr Datenverlust zu erleiden. Ist der Kontrast im Nummernschild einmal so klein, dass er nicht mehr ausreicht, um das Kennzeichen in dem rekombinierten Bild gleichzeitig mit dem Fahrer sichtbar zu machen, so stellt der lineare Bereich bei der Logarithmierung sicher, dass keine wichtigen Informationen verloren gehen. Dabei muss berücksichtigt werden, dass nicht alle Aufnahmeverfahren, die zur Zeit im Rahmen der Verkehrsphotographie zur Verfügung stehen, getestet wurden.

Für die Zukunft kann man sich vorstellen, das Ergebnis noch weiter zu verbessern, in dem z.B. die Lesbarkeit des Kennzeichen mit morphologischen Filtern hervorgehoben wird. Dies soll aber nicht mehr Bestandteil dieser Arbeit sein. Auch der Zweck, bzw. der Nutzen ist fraglich, weil die vorgestellte Arbeit schon sehr gute Ergebnisse liefert und unter Umständen der Rechenaufwand dadurch zu hoch wird.

Auch die Behandlung von Farbbildern kann man sich mit dem Verfahren der nichtlinearen Verarbeitung in einer Laplacepyramide in Zukunft vorstellen.

5 Anhang

5.1 Literaturverzeichnis

- [1] Schröder, „Technische Fotografie“, (Vogel), 1.Auflage, 1981
ISBN - 3-8023-0144-1
- [2] Zamperoni, „Methoden der dig. Bildsignalverarbeitung“, (Vieweg), 1989
ISBN - 3528-03365-7, 1989
- [3] Burt & Adelson, „The Laplacian Pyramid as a Compact Image Code“, 1983;
IEEE Transaction on communications, Vol. com-31, No.4, pp. 532-540
- [4] Jähne, „Digitale Bildverarbeitung“, (Springer), 3. Auflage, ISBN - 540-56926-X, 1993
- [5] Jähne, „Digitale Bildverarbeitung“, (Springer), 4. Auflage, ISBN - 540-61379-X, 1997
- [6] http://www.uni-kiel.de/ewf/geographie/lehre/fe-vorlesung_ss99/m2/sld016.htm Dez, 2002
- [7] http://www.ps-explore.de/html/histogramm_0.html Dez, 2002
- [8] <http://www.fo.fh-koeln.de/dozenten/dkunz/skripte/bv/> Dez, 2002
- [9] <http://www.telematik.informatik.uni-karlsruhe.de/lehre/WS9900/vorlesungen/Telematik/K03-Nachrichtentechnische-Grundlagen/sld075.htm> Dez, 2002

5.5 Verwendete Bildformate

1.) *.biu Format :

Die Bilddaten befinden sich unverschlüsselt und unkomprimiert in Binärform. Die Größe des Bildes steht am Anfang der Datei. Die ersten 2 Bytes beinhalten die Größe der X-Dimension und die darauf folgen 2 Bytes die Größe der Y-Dimension. Anschließend folgen die Bilddaten. Auch hier stehen immer 2 Bytes für ein Pixel (16bit).

X-Dimension	Y-Dimension	1.Pixel	2.Pixel
00 02	00 02	85 00	A1 00
$0 \cdot 16^0 = 0$	$0 \cdot 16^0 = 0$	$5 \cdot 16^0 = 5$	$1 \cdot 16^0 = 1$
$+ 0 \cdot 16^1 = 0$	$+ 0 \cdot 16^1 = 0$	$+ 8 \cdot 16^1 = 128$	$+ 10 \cdot 16^1 = 160$
$+ 2 \cdot 16^2 = 512$	$+ 2 \cdot 16^2 = 512$	$+ 0 \cdot 16^2 = 0$	$+ 0 \cdot 16^2 = 0$
$+ 0 \cdot 16^3 = 0$	$+ 0 \cdot 16^3 = 0$	$+ 0 \cdot 16^3 = 0$	$+ 0 \cdot 16^3 = 0$
<u>512</u>	<u>512</u>	<u>133</u>	<u>161</u>

Abb. 5.5.1 : Aufbau *.biu Format im hexdezimal System

2.) *.raw Format :

Ebenso wie das *.biu Format liegen die Bilddaten in binärer Form vor. Der einzige Unterschied beruht auf dem erweiterten Header:

Byte 01 - 04 : Fotonummer; Byte 05 - 08 Bildbreite in Pixel

Byte 09 - 12 : Bildhöhe in Pixel; Byte 09 - 12 : bits/pixel

Byte 17 - 20 : Farbe, 1 für Schwarz

3.) *.bid Format :

Aufgebaut wie das *.biu Format, nur sind die Pixel in einem Byte (8bit) angegeben.

5.6 Testbilder

Alle folgenden Bilder sind nach dem Parametersatz aus 3.4 gefiltert, über die Funktion „Auslesen_8_bit“ als *.bid Datei ausgegeben und in Bild.exe als Tiff Bild abgespeichert.



Originalbild, Darstellung untere 8 bit



Originalbild, Darstellung obere 8 bit

055 km/h R DIR h min s DAT. CODE FOTO TRAFFIPAX
055 1 ↓↓ 14:43:58 25.09.01 022 3334 154 RADAR





Originalbild, Darstellung untere 8 bit



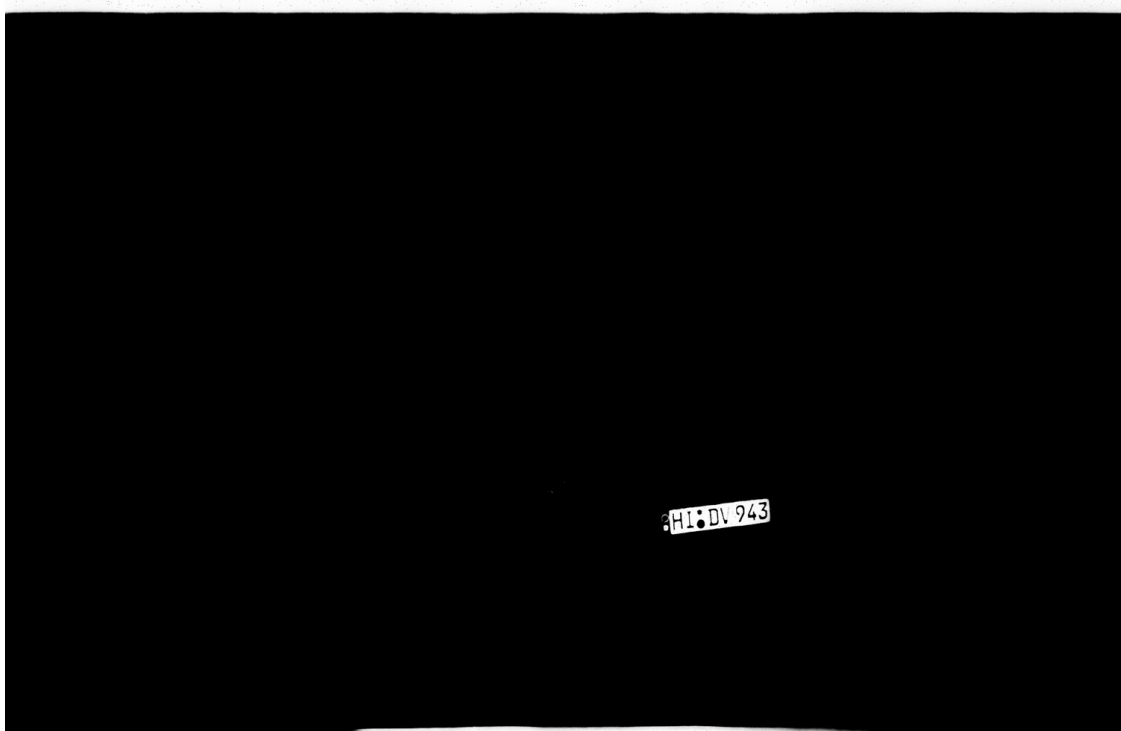
Originalbild, Darstellung obere 8 bit

051 km/h R DIR h min s DAT . CODE FOTO TRAFFIPAX
051 1 ↓↓ 14:43:16 25.09.01 022 3334 153 RADAR





Originalbild, Darstellung untere 8 bit



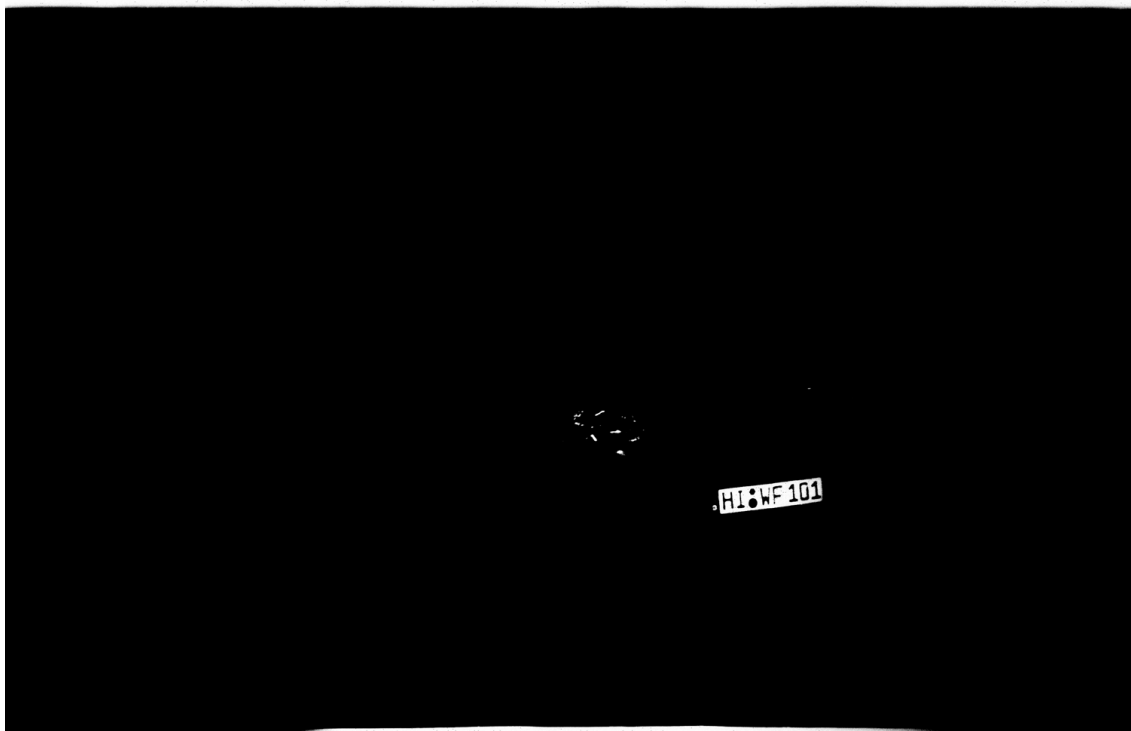
Originalbild, Darstellung obere 8 bit

062 km/h R DIR h min s DAT CODE FOTO TRAFFIPAX
062 1 JJ 14:41:21 25.09.01 022 3334 149 RADAR





Originalbild, Darstellung untere 8 bit



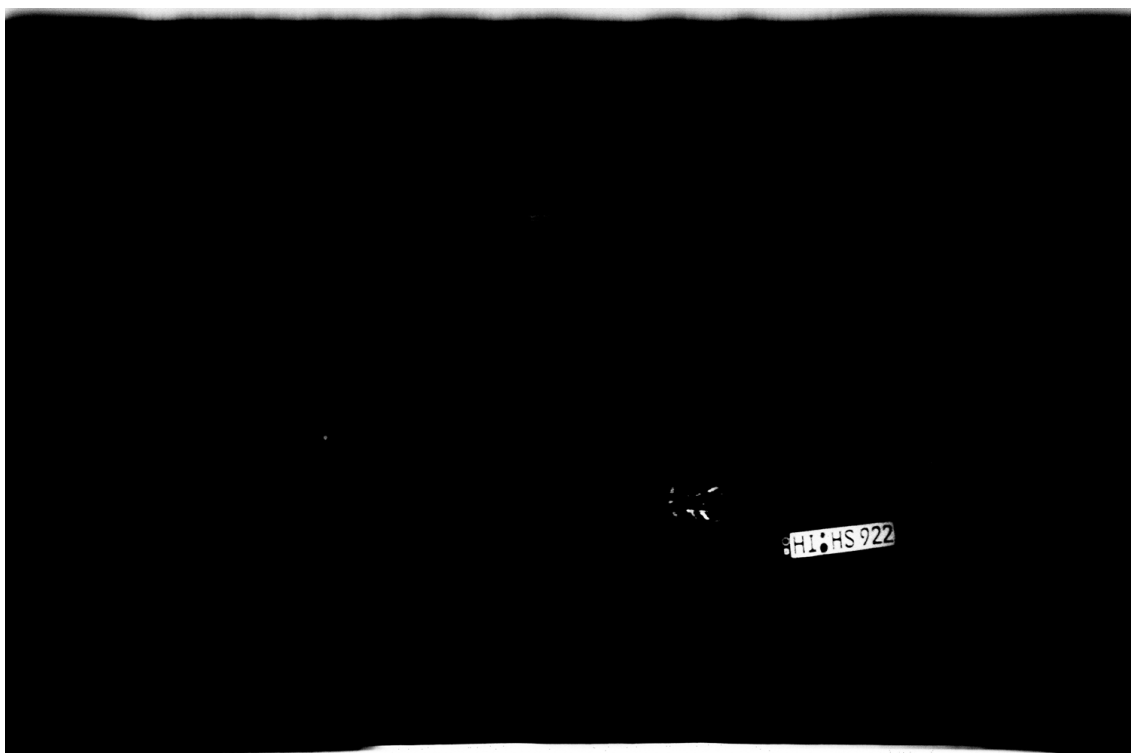
Originalbild, Darstellung obere 8 bit

055 km/h R DIR h min s DAT. CODE FOTO TRAFFIPAX
055 1 ↓↓ 14:44:52 25.09.01 022 3334 155 RADAR





Originalbild, Darstellung untere 8 bit



Originalbild, Darstellung obere 8 bit

063 km/h R DIR h min s DAT. CODE FOTO TRAFFIPAX
063 1 ↓↓ 14:52:49 25.09.01 022 3334 170 RADAR





Originalbild



gefiltertes Bild

5.6 Eidesstattliche Erklärung

Ich versichere hiermit, die vorgelegte Arbeit in dem gemeldeten Zeitraum ohne fremde Hilfe verfasst und mich keiner anderen als der angegebenen Hilfsmittel und Quellen bedient zu haben.

Köln, den 20.12.2002

Unterschrift

5.7 Sperrvermerk

Die Einsicht in die vorgelegte Arbeit ist bis zum 20.12.2004 gesperrt.

5.8 Weitergabeerklärung

Ich erkläre hiermit mein Einverständnis, dass das vorliegende Exemplar meiner Diplomarbeit oder eine Kopie hiervon für wissenschaftliche Zwecke verwendet werden darf.

Köln, den 20.12.2002

Unterschrift